Risks of Using AP Locations Discovered Through War Driving

Minkyong Kim, Jeffrey J. Fielding, and David Kotz

Department of Computer Science
Dartmouth College

Abstract. Many pervasive-computing applications depend on knowledge of user location. Because most current location-sensing techniques work only either indoors or outdoors, researchers have started using 802.11 beacon frames from access points (APs) to provide broader coverage. To use 802.11 beacons, they need to know AP locations. Because the actual locations are often unavailable, they use estimated locations from *war driving*. But these estimated locations may be different from actual locations. In this paper, we analyzed the errors in these estimates and the effect of these errors on other applications that depend on them. We found that the estimated AP locations have a median error of 32 meters. We considered the error in tracking user positions both indoors and outdoors. Using actual AP locations, we could improve the accuracy as much as 70% for indoors and 59% for outdoors. We also analyzed the effect of using estimated AP locations in computing AP coverage range and estimating interference among APs. The coverage range appeared to be shorter and the interference appeared to be more severe than in reality.

1 Introduction

Pervasive computing applications often need to know the location of users. This location information should be available anywhere, both indoors and outdoors. While some location sensing techniques, such as Cricket [11] and Bat [5], provide high accuracy, they are mostly limited to indoor usage. On the other hand, satellite navigation systems like GPS [4] are useful for outdoor navigation, but they do not work well in urban settings due to the “urban canyon” effect.

To address these limits, researchers have started using 802.11 beacon frames from access points (APs) to locate wireless network users. Intel’s Place Lab [3, 8] provides software that can track users both indoors and outdoors. Skyhook Wireless [13] provides a similar commercial solution for locating Wi-Fi users. These approaches require knowledge of the (actual or estimated) location of APs. In addition to user-location tracking, researchers also use the location of APs to analyze wireless network characteristics such as the coverage range of APs or interference among APs.

Although we may be able to get the actual location of APs for managed networks, it is almost impossible to get the actual location of unmanaged networks. Thus, researchers [3, 1] recently started using the AP locations estimated through *war driving*. War driving is the process of collecting Wi-Fi beacons by driving or walking through a town, to discover and map the location of APs [7]. Because war driving is easy and

K.P. Fishkin et al. (Eds.): PERVASIVE 2006, LNCS 3968, pp. 67–82, 2006.

can be performed by anybody with a wireless card, a GPS receiver, and war-driving software, it is an effective way of collecting AP location information. The AP locations determined by war driving, however, are estimates rather than actual locations. Thus, it is important to understand the errors in these estimates and the effect of these errors on other applications that depend on these estimated AP locations.

The main goal of this paper is to understand the effect of using AP locations estimated through war driving. We do not want to discourage people from using the estimated AP locations, but rather we want to encourage them to use the data with an appropriate caution. Our focus is on comparing various results using estimated AP locations against those using actual AP locations. We explored the error in estimated AP locations. The median error in estimated AP locations was 32 meters. We considered the error in tracking user positions both indoors and outdoors. Using actual AP locations, we could improve the accuracy of user location estimates as much as 59% outdoors and 70% indoors. We also analyzed the effect of using estimated AP locations in analyzing AP coverage ranges and estimating the inferences among APs. The coverage range appeared to be shorter when estimated AP locations are used and the interference among APs appeared to be more severe.

2 Related Work

Many wireless network users use war-driving data to learn the location of APs for (free) Wi-Fi connectivity. There are several Internet Web sites, including WiFiMaps.com [15], that collect and provide this information. People discover Wi-Fi hotspots through war driving and upload their data to these sites. As the main goal of these Web sites is to discover available Wi-Fi connectivity, it is not important to accurately estimate the location of APs, but this may not be the case for other applications.

Although the accuracy of AP location estimates can be improved with additional hardware, such as directional antennas [12], it is often more time consuming to collect data using extra hardware and this hardware is not commonly available among typical Wi-Fi users. War driving without extra hardware seems to be an easy and convenient way to collect AP locations for larger areas, although its estimates may be inaccurate.

There are many applications that need to know the accurate location of APs. *Localization* is the process of determining the physical location of a user. Localization techniques that use Wi-Fi beacons depend on accurate information about the location of APs. Place Lab [3] uses the AP location estimates from war driving to track a user's location. Other localization techniques [2] assume that the locations of *reference points* are known without specifying methods to discover their locations. Just as in Place Lab, one could use the location estimates from war driving for these approaches. Besides localization, researchers have started using AP location estimates to study AP deployment characteristics such as AP density and interference among APs [1]. Unlike the original motivation for war driving, which is finding Wi-Fi hotspots, these applications are highly affected by the accuracy of AP locations. Thus, it is critical to analyze the accuracy of estimated AP locations and understand their impact on the applications that depend on them. The only previous work that analyzed the accuracy of AP locations estimated through war driving is not comprehensive; it considered the location of only five APs [9]. To the best of our knowledge, ours is the first research study to analyze in a large scale the accuracy and the impact of using AP location estimates.

3 Methodology

As researchers have started using data collected by war driving for applications such as localization, it is important to understand the errors in war-driving data. Given the actual AP locations on our college campus, we performed war driving on the campus and obtained the estimated AP locations.

We believe that the Dartmouth college campus is an ideal place to perform this study. First, Dartmouth has wireless coverage almost everywhere on the campus. Second, all APs on the campus are centrally administrated. Thus, it is relatively straight-forward to obtain information about these APs. Third, information about the location of APs is up-to-date since we have recently replaced all of our APs and recorded detailed location data.

In the following sections, we describe the process of mapping APs on our campus map, war driving on the campus, and the algorithms from Place Lab that we used to estimate AP locations and to track user positions.

3.1 Actual AP Locations

To understand the effect of using AP location estimates, we first need to obtain the actual AP locations to serve as the 'truth'. We were lucky to have access to the actual AP locations on our college campus. Our network administrators keep records of the location of APs on floor plans of campus buildings. Using these floor plans, we determined the precise location of APs on the campus map. These locations serve as the *actual* locations. In this way, we mapped 927 APs. Out of 927, 44 APs are dedicated to air monitoring and the rest are regular APs. The air monitors collect network statistics and work only in a passive mode, not sending out any signals. Except seven APs that support only 802.11g, all APs support both 802.11a and 802.11g. While 100% of the APs on our campus support 802.11g, this ratio is much lower for observed unplanned networks in Pittsburgh: 20% supporting 802.11g and the rest supporting only 802.11b [1].

3.2 War Driving

To understand the effect of using estimated AP locations through war driving, we drove and walked around our campus. We used a Linux laptop and a Cisco *Aironet 350* wireless card, which supports 802.11b. The laptop ran the *Place Lab stumbler 2.0* to collect beacons from APs. We also carried a GPS device, Garmin *etrex*, attached to the laptop.

We drove around the campus with these devices at a speed of 10 miles/hour or less to allow the wireless card enough time to pick up beacons. Our war driving lasted about 80 minutes. Since we could not drive close to many buildings, we decided to augment the war-driving data with *war walking*.

We walked around the main parts of the campus to cover the areas that cars cannot reach. We collected war-walking data for about 200 minutes. Because both war driving and war walking use GPS readings to locate the position of the recorder, we had to stay outdoors. To get signals from as many APs as possible and also not to bias the AP-location estimates towards one direction, we walked *around* each building and tried to stay close to it as long as we had GPS signal reception. Unfortunately, we often encountered obstructions—such as trees, outside structures, and construction vehicles—that prevented us from walking close to buildings.

3.3 Algorithms

Intel's Place Lab project [10] is well-known for using war driving data to locate APs and perform localization by detecting Wi-Fi beacons from APs. We use the software provided by Place Lab to estimate AP and user locations.

To estimate AP locations from war driving and war walking, we looked into three positioning algorithms: centroid, weighted centroid, and particle filters. Given n location measurements, the geometric centroid $\bar{\mathbf{x}}$ is defined as $\bar{\mathbf{x}} = \sum_{i=1}^{n} \mathbf{x}_i / n$ where $\mathbf{x}_i$ is location of the ith measurement. This simple centroid does not consider the signal strength of beacons. The weighted centroid considers signal strength received during the scan. During our war walking and war driving, we observed values between -123 dBm and -25 dBm. These values are linearly mapped to values between 0 and 100 and then used as the weights for the weighted-centroid algorithm. The particle filter [3] is based on Bayes' theory. To estimate an AP's location, it uses a sensor model that assigns probabilities to particles based on the observed signal strength and the distance from the particle to the observer. The default motion model is null since APs do not move.

To estimate user position, we use a particle filter with a sensor model that describes the likelihood of observing a set of APs with their received signal strengths given the particle's distance to each AP. The default motion model moves particles random distances in random directions. Details on particle filters can be found in Hightower and Borriello's paper [6].

4 Understanding War-Driving Data

Our main goal is to understand the effect of using estimated AP locations rather than actual locations on user-location tracking and wireless network characterizations. More specifically, we explore following questions:

- How effective is war driving or war walking in discovering APs?
- How well can we estimate the location of APs by war driving or war walking?
- How well can we track user positions outdoors?
- How well can we track user positions indoors?
- What is the effect of using estimated AP locations on analyzing AP coverage range?
- What is the effect of inaccuracy in AP locations on analyzing AP interference?

4.1 Effectiveness of War Driving

In this section, we consider the effectiveness of war driving or war walking in discovering APs. Excluding 44 air monitors, we know the actual location of 883 APs deployed on our campus. Out of 883 APs, we detected only 334 APs during war driving, and detected an additional 187 APs through war walking. This makes the AP detection rate 38% for war driving and 59% for the combination of war driving and war walking. We also detected 172 APs whose actual locations are unknown. We exclude these APs in our analysis of errors in AP locations since we do not have the ground truth, but later analyze their effect on estimating user locations (see Fig. 7 and 9). Table 1 summarizes the number of APs detected during war driving and war walking.

Figure 1 shows the estimated location of APs on the Dartmouth campus map. We see that the weighted-centroid algorithm estimated APs to be close to the war-driving or war-walking tracks; estimated AP locations are often on the tracks recorded by the GPS device. This tendency is especially strong around the edge of the campus where there are fewer roads. One side effect of this outcome is that APs appear to be close to each other. We later consider the consequences of this in Section 4.6.

Out of 883 APs, we only detected 521 APs during war driving and war walking. Most of the 362 undetected APs are in the outer region of the campus where we did not

Table 1. APs detected during war driving and war walking. Note that the total is smaller than the sum of war driving and war walking because many APs are detected during both.

	APs w/ known location	APs w/ unknown location
Driving	334	56
Walking	384	155
Total	521	172

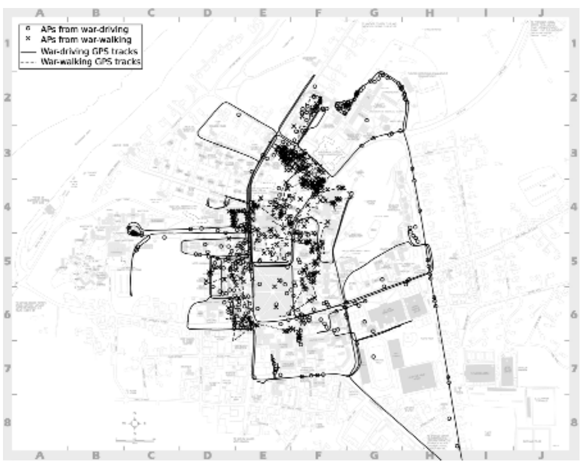


Fig. 1. Estimated AP locations on campus map. '○' marks denote 521 APs detected during war driving and 'x' marks shows additional APs detected by war walking. Lines show war-driving and war-walking paths recorded by a GPS device.

do war walking (e.g., the west end of the campus). But, some APs are actually in areas that we walked around; these APs are inside large buildings (such as the main library in the center of the campus) and mostly in the basement or on higher floors of buildings.[1] Being outdoors on ground-level apparently prevented us from detecting signals from APs in basements and on upper floors.

One of the important characteristics in understanding AP deployment is the density of APs. The size of our main campus is roughly 1 km^2. Using this size, the density using the actual number of APs is 927/km^2, while those based on the APs discovered by war driving and war driving with walking (*driving-walking*) are 334/km^2 and 521/km^2, respectively. Cheng et al. [3] reported the density of three neighborhoods in the Seattle area. Our density of 927/km^2 is close to those of the downtown Seattle (1030) and Ravenna (1000), while it is higher than that of Kirkland (130). Note that in computing density, we considered only the APs whose locations are known; if we include 172 APs with unknown locations, we get the density of 1099/km^2.

We also present the number of APs detected at each scan by the Place Lab stumbler; the stumbler scanned every two seconds. Figure 2 shows the cumulative fraction of scans as a function of the number of APs for each scan. It includes the result for war driving, war walking, and the two combined. War walking detected more APs than war driving: The averages are 11.5 and 6.1 for war walking and war driving, respectively. Since we did not, or could not, take exactly the same paths for war driving and war walking, it may not be fair to directly compare these two averages, but war walking in general seems to be more effective in detecting beacons than war driving. We expect that this is because war walking is slower and its paths are closer to buildings where APs are located. The average for combined is 10.0 APs per scan; this average is much higher than the average reported by Cheng et al. [3] for three neighborhoods in the Seattle area—2.66, 2.56 and 1.41—although the density of APs in two studies are similar. This is mostly due to the fact that we augmented war-driving data by war walking, while Cheng et al. collected traces only by war driving.

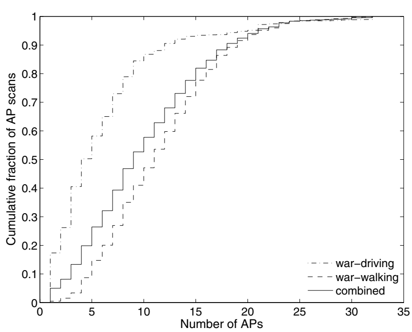


Fig. 2. Number of APs detected per scan

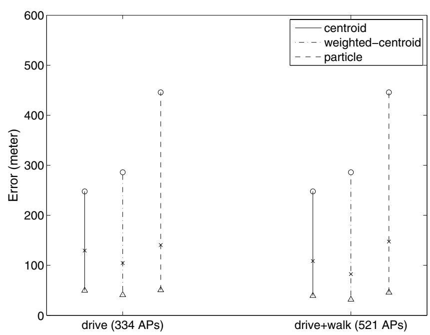


Fig. 3. Error in AP locations. △, ×, and ○ marks show the value for 50%, 95%, and 100%, respectively

[1] The highest building on our campus has six floors above the ground level.

4.2 AP Locations

In this section, we consider the error in AP-location estimates from war driving and war walking using known AP locations on our campus. Figure 3 shows the error in AP locations using three positioning algorithms. We consider two sets of APs: one discovered by war driving only and the other discovered by either war driving or war walking. We see that war walking helped reduce the error for both simple-centroid and weighted-centroid algorithms; we are not sure why this is not the case for the particle filter. For both sets of APs, the weighted centroid outperformed both the simple centroid and the particle filter. Its median error using war-driving data was 40.8 meters, while the error using both war-driving and war-walking data was 31.6 meters. In the remainder of the paper, we only consider the AP location estimates generated by the weighted centroid.

4.3 User Location: Outdoor

To understand the effect of using estimated AP locations to estimate user position, we walked along four, mostly non-overlapping, paths. Together, they cover the central part of the campus. Each walk lasted around 10 minutes, including a one-minute pause at a location. During these four outdoor walks, we detected 8 additional APs. We did not include them in our war-walking AP set because adding the traces from test walks affects the results. Figure 4 shows these four walks on the campus map.

We used Place Lab's particle filter [3] to estimate user position using the beacon data collected during the four walks. We estimated user paths using three sets of the AP locations: actual, war driving only, and war driving and war walking combined.

Figure 5 shows the paths for the four test walks. For each walk, we plotted the GPS track and the estimated paths using the three sets of AP locations. The circles on the GPS track denote the location of the one-minute pauses. When looking at Walk 3, the estimated tracks using AP locations from war driving and driving-walking were particularly inaccurate. This inaccuracy is due to a big open area, which does not contain any APs but is covered by several powerful APs around it. For Walk 4, estimated paths

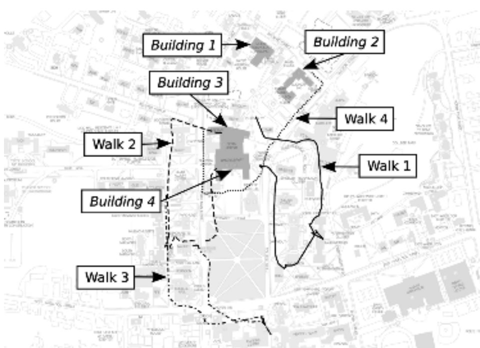
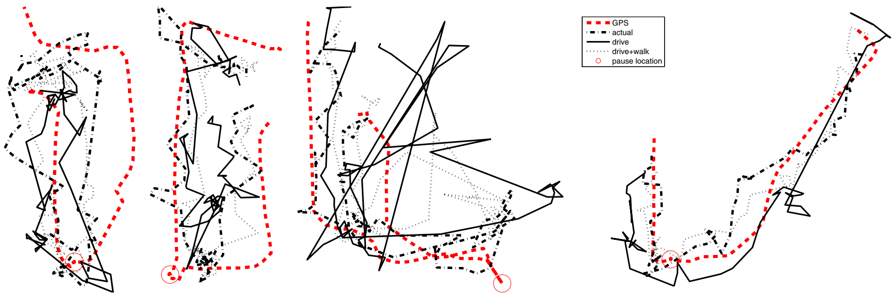


Fig. 4. Test walks on campus map. This figure depicts GPS tracks of four outdoor test walks and the locations of four buildings where indoor test walks were performed.



(a) Walk 1 (b) Walk 2 (c) Walk 3 (d) Walk 4

Fig. 5. User location

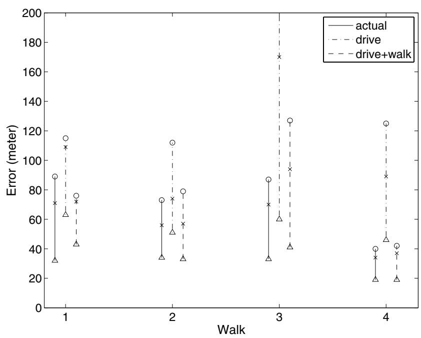


Fig. 6. Error in outdoor user position (△:50%, ×:95%, ○:100%). The 95% and 100% values for Walk 3's drive group are 231 and 250, respectively.

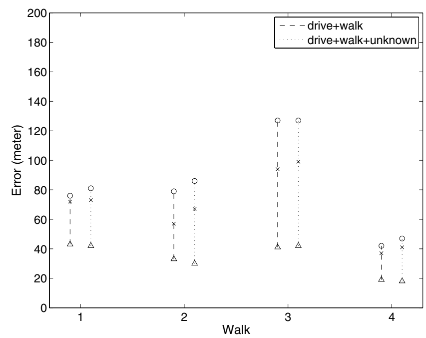


Fig. 7. Error in outdoor user positions with unknown APs (△:50%, ×:95%, ○:100%)

were close to its GPS track because this walk was through an area with dense APs. For all four walks, the estimated paths converged with GPS tracks near the pause locations, presumably because the estimator corrected the user location as it detected beacons from more APs located near the pause location.

Figure 6 shows the error in user location estimates. The GPS tracks again served as the ground truth. The error is the difference between the GPS tracks and the estimated paths, computed every 20 seconds. For each walk, errors with actual, war-driving, and war-driving with war-walking AP locations are shown. The errors using actual AP locations were much smaller than those using the war-driving AP locations. Compared to errors using driving-walking, the errors of the actual set are smaller for Walk 3, and about the same for the other walks. The reason for this closeness is because we walked along similar paths during the war walking. In summary, the median error in four walks using actual AP locations ranged 19–34 m, that for war driving ranged 46–63 m, and that for driving-walking ranged 19–43 m.

Although Figure 6 clearly shows that actual AP locations outperformed war driving, one might wonder whether it is due to the fact that we know the location of a bigger number of APs (883). To factor out this problem, we ignored the actual location of APs

Table 2. Improvement in outdoor user-location estimates by using actual AP locations instead of estimated AP locations. We used a subset of actual locations of APs that were detected during war driving and driving-walking. The median of normalized improvement is depicted.

Walk number	1	2	3	4
Outdoor: drive	33%	31%	17%	59%
Outdoor: drive-walk	13%	5%	28%	14%

that were not detected during war driving and war walking. Thus, we only used the same set of APs that were detected during war driving; we did the same for driving-walking. Table 2 shows the result. Each number is the median of the normalized improvement for every user position: $median(\frac{(e_i - a_i)}{e_i} \times 100\%)$ where e_i and a_i are the errors using estimated and actual AP locations, respectively, for the ith position. We used 334 actual AP locations for war driving, and 521 locations for driving-walking. When we used the actual AP locations, the accuracy improved for both war driving and driving-walking. The median improvement ranged 5%–59%. Having accurate AP locations is important to estimate user position correctly.

During war driving and walking, we discovered 172 APs whose actual locations are unknown. We considered whether using these extra APs reduces the user location error. Figure 7 shows the error for the two sets of APs: drive-walk and drive-walk with unknown APs. Note that all these AP locations are estimates. The result shows that using extra APs did not make much difference; for all four walks, it even increased the 95% error values.

4.4 User Location: Indoor

The main benefit of using Place Lab over a GPS device is that it is usable where a GPS signal is not available. A Place Lab paper by Cheng et al. [3] contains a simple evaluation of indoor accuracy of user positions. They visited nine indoor locations and found that the location error ranges from 9 to 98 meters. In this section, we further explore the accuracy of Place Lab in tracking user positions indoors.

To estimate our position indoors, we marked points on the building's floor plan as we walked. While the latest Place Lab source includes a stumbler using this method, we had trouble getting it to work on our stumbler laptop. In addition, many of the floor plans were oriented at odd angles, which Place Lab's *mapwad* format[2] does not support. We wrote a new stumbler to map the floor plans to the campus map using conversions we derived earlier as we mapped the known AP locations on the campus map.

We chose four buildings, most of whose APs were detected during either war driving or war walking; we in fact walked around each of these four building during war-walking. Figure 4 shows the location of these buildings on the campus map. We walked inside of these four buildings, covering three or more floors within each building. The duration of our indoor walks ranged from 6 to 14 minutes; this duration does not include time that we took to move to the next floor. Table 3 shows the duration and the floors that we walked within each building.

[2] A format that includes maps and sets of places [10].

Table 3. Indoor walks. This table shows the floors that we walked within each building.

Building	Duration (minute)	Floors
1	14	Basement, floor 1, 2, 3 and 4
2	9	Basement, floor 1 and 2
3	7	Floor 1, 3 and 4
4	6	Basement, floor 1 and 3

Table 4. Improvement in indoor user-location estimates. We used a subset of actual locations of APs that were detected either during war driving or war walking.

Building number	1	2	3	4
Indoor: drive-walk	14%	53%	17%	70%

During our indoor test walks, we detected an additional 24 APs, many more than the 8 discovered in our outdoor walks. This result is not surprising since we stayed only outdoors during war driving and war walking. We can imagine using our software extension to do war walking indoors to augment the data collected by outdoor war driving and walking. But, indoor war walking may not be possible in other situations since it requires physical access to buildings and digitized floor plans of those buildings.

Figure 8 shows the error in user positions using the actual AP locations and war driving with walking. We could not use the AP locations estimated from war driving only because we did not have enough APs to make reasonable user-location estimates for indoor walks. For example, through war driving, we did not discover any APs in Building 1. The error in user position was computed every 10 seconds.

Compared to outdoor walks, indoor walks had smaller absolute errors. This difference is partly because each indoor walk covered a smaller area than the outdoor walks. The median error using actual AP locations ranged from 7 to 11 meters, while that using estimated AP locations ranged from 15 to 30 meters. On the other hand, the relative difference between using actual and estimated AP locations was bigger than the difference for outdoor walks because we could not do war driving or war walking inside of buildings and the estimated AP locations were often close to war-driving or war-walking paths. Note that Building 1 had some big errors; its maximum error was 91 meters. These large errors were due to the walk in the basement, which is under the ground level without any windows; the other two buildings (2 and 4), in which we also walked in the basement, are half underground with windows.

As we did for the outdoor test walks, we computed the median relative improvement for indoor walks. We used the actual location of the same set of APs that were detected during war driving and walking to compute the user-position errors. We then computed improvements in user position errors using these actual AP locations. Table 4 shows the median relative improvement, which ranged 14–70%. The average indoor improvement for four buildings (38.5%) was greater than that for outdoor walks with driving-walking (15.0%).

We now consider the effect of using the extra 172 APs in estimating indoor user location. Figure 9 shows the error for both driving-walking and driving-walking with

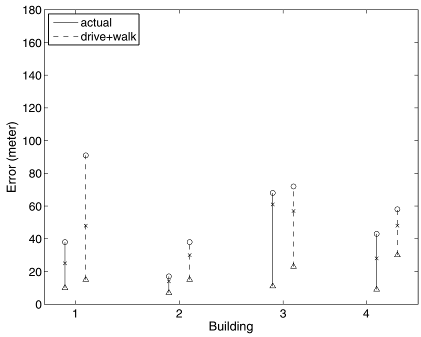


Fig. 8. Error in indoor user location, computed every 10 seconds. (△:50%, ×:95%, ○:100%).

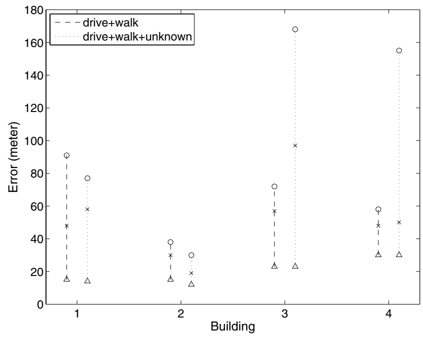


Fig. 9. Error in indoor user location with unknown APs. (△:50%, ×:95%, ○:100%).

unknown APs. Using extra APs did not make much difference in median errors; it reduced the 95% error for Building 2, but increased it for the other three buildings. Although there was a much larger worst-case error in two buildings, these cases represent outliers. In summary, for both outdoor and indoors, using extra APs did not make much difference.

For context-aware applications such as a shopping assistant [14], it is important to know whether a user is inside of a building or outdoors. For example, if a user is inside of a grocery store, his context-aware application may pull up the list of items that are running short at home. On the other hand, if a user is passing by a grocery, the application should not pull up the items but just remind the user that he may need to do some grocery shopping.

Figure 10 shows the percentage of estimates for which the particle filter correctly estimated user position to be inside using three different sets of APs: actual location of 883 APs, actual location of 521 APs, and estimated location of 521 APs. (During indoor walks, the user was inside 100% of time.) On average for four buildings, the filter was correct 76% and 71% of time using actual location of 883 APs and 521 APs, respectively, but it was correct only 42% of time using estimated AP locations. Note that using the actual location of 521 APs produced smaller errors than using 883 APs for Building 1; we do not yet have a clear explanation for this result.

4.5 Maximum Signal Coverage

On our campus, we have three models of Aruba APs: 52, 60, and 72. These Aruba APs adjust their power level dynamically and some of them are deployed with special antennas to cover larger areas. Furthermore, signal propagation in a complex environment is difficult to predict. Thus, it is hard to specify the signal coverage range. Instead, we computed it from empirical data gathered while war driving and walking.

For each AP, we computed the distance from the known AP location to the farthest point where each AP was detected during war driving or war walking. Figure 11 shows the cumulative fraction of APs as a function of the maximum signal range. To see the effect of using inaccurate AP locations, we also included the maximum signal range computed from AP locations estimated from war driving and driving-walking. Note

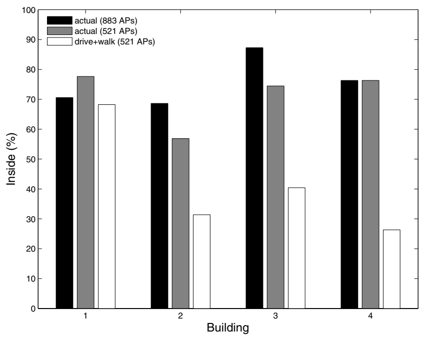


Fig. 10. Inside buildings. This figure shows the percentage of estimates for which the algorithm correctly indicated a user location as inside of each building.

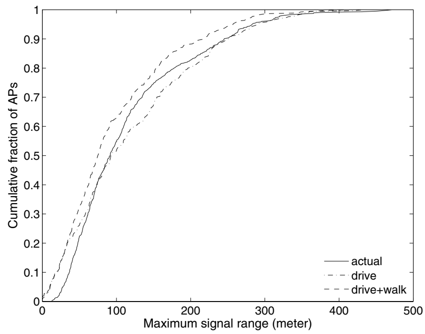


Fig. 11. Maximum signal coverage. The x-axis shows the maximum signal range observed during war driving and war walking. The y-axis shows the cumulative fraction of APs.

that because we used the data collected by war driving and war walking, the recorded maximum range is only an approximation. (Ideally, we should circle around each AP, increasing the radius for each round until we do not hear the signal.) Nonetheless, the median using actual AP locations was exactly equal to the commonly believed range of 300 feet (91.4 m). The maximum observed range of all APs was 470.8 m.

When we used AP locations estimated from war driving or driving-walking, we found more APs with small ranges than when the actual AP locations were used. This results from the tendency of war driving and war walking to estimate APs to be close to where beacons were detected, away from their true locations and closer to roads. The medians were 93.2 and 74.9 meters for war driving and driving-walking, respectively. The maximums were 431.8 and 413.0 meters.

4.6 AP Interference

Although there needs to be some overlap in AP coverage areas to have seamless wireless connectivity, overlaps among too many APs reduce the effective throughput. We computed the interference using AP locations with the 50-meter range used by Akella et al. [1], who focused on characterizing wireless networks that are unplanned and unmanaged based on the assumption that these networks suffer from higher interference than planned-managed networks. Here, we extract the same set of characteristics from our planned-managed campus network.

The data in preceding sections of the paper are based on a campus-wide deployment of 927 APs, as described in Section 3.1. Two months later, we set out to explore interference and the network had grown to 1042 APs, out of which 47 were air monitors. In analyzing interference, we consider only the 995 regular APs.

Using actual AP locations: To analyze interference among APs, Akella et al. presented the degree of each AP, where degree is defined as the number of other APs in interfering range. Figure 12 shows the degrees of 995 APs computed using their actual locations and the 50 m range assumption. Out of 995 APs, 976 APs had 3 or more neighbors.

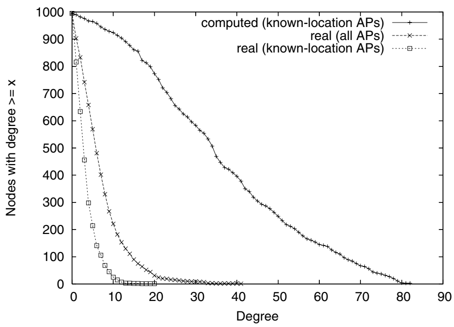


Fig. 12. AP degrees. 'computed' denotes degrees computed with the 50 m interference assumption. Two 'real' lines present degrees observed by APs: one with all APs and the other including only the APs with known locations.

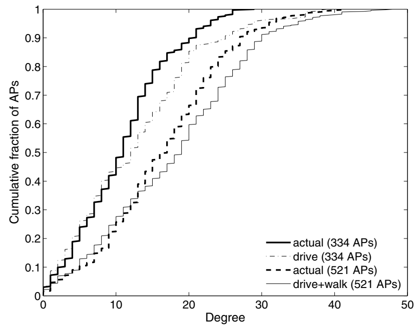


Fig. 13. AP interference. This figure compares the degrees computed using estimated and actual AP locations. For both war driving and driving-walking, the degree using actual locations is smaller than that using estimates.

This ratio of APs with 3 or more neighbors (98%) is actually higher than those for all six cities that Akella et al. reported (approximately 25%–80%). This result suggests that these nodes interfere with at least one other node since only three of the 802.11b channels do not interfere much. The maximum AP degree for our campus was 82. This number is high in the range (20–85) that Akella et al. reported. In short, we found that the computed interference was actually more severe in our planned campus network than the unplanned networks considered by Akella et al.

Figure 12 also includes the real degrees obtained from APs. Each Aruba AP checks periodically which beacons it can detect on every channel and the master switch aggregates this information from APs. Our 995 APs detected 1234 APs, including third-party APs whose locations are unknown. We report two sets of data: one counting only the 995 APs that we know the location of, and the other counting all 1234 APs that our 995 APs detected. The former is included to compare against the computed degree, while the latter presents real observed values. We can see clearly that the real degrees were much smaller than ones computed using the 50-meter interference assumption. The average degree from real data considering all APs is 6.6 and that considering only known-location APs is 2.9. In contrast, the average degree with the 50-meter assumption is 35.0 APs.

The main reason that the computed interference degree is much larger than the real observed one is the fact that our campus wireless network is a planned-managed network and the power levels on APs are adjusted to minimize interference. We expect that the difference will be smaller for unplanned-unmanaged networks [1]. Another reason is that, in reality, obstructions can prevent APs from hearing each other even they are located close to each other.

The interference degree may in fact be even less than the values depicted as 'real' in Figure 12 because multiple channels are used for APs and APs on different channels do not interfere much. The 802.11g APs on our campus are evenly divided into four channels: 1, 4, 8, and 11. We consider the interference degrees observed within each channel, that is, the number of APs whose beacons can be heard on the same channel.

Table 5. AP degrees for different channels of 802.11g

Channel	Number of APs	Average degree	Max degree
1	294	2.3	13
4	229	1.1	8
8	202	1.2	8
11	270	2.6	12

Table 6. Summary. The column 'Actual' shows the values using the actual AP locations. 'D' and 'D+W' denote the results using the estimated location of APs discovered through driving and driving-walking, respectively.

Analysis	Actual	D	D+W
Effectiveness of war driving: War walking was more effective than war driving at detecting APs.			
Percentage of APs detected (%)		38	59
Average number of APs for each scan (* war walking only)		6.1	11.5*
AP location: APs often appear to be closer to roads, or to each other, than in reality.			
Median error in AP location estimates (meter)		40.8	31.6
Outdoor user location: Having accurate AP locations is important to estimate user position correctly.			
Median error over four walks (meter)	19-34	46-63	19-43
Median improvement using the actual AP locations (%)		17-59	5-28
Indoor user location: The improvement obtained using actual AP locations was greater for indoor walks than outdoor walks.			
Median error for four walks (meter)	7-11		15-30
Median improvement using the actual AP locations (%)			14-70
Inside of buildings: Using war driving and walking data poorly estimated whether a user is inside.			
Percentage of correct estimation (%)	76		42
Signal range: Using estimated AP locations, some APs appeared to have shorter signal ranges than when actual locations were used.			
Median signal range (meter)	91.4	93.2	74.9
AP interference: The computed degree using the actual locations (35.0) was overestimated compared to the real interference (2.9). Using the estimated AP locations made the degree even bigger.			
Median degree difference using estimated and actual locations		2	1

Table 5 presents the number of APs on each channel, the average degree, and the maximum degree as observed by the APs. This result includes APs whose locations are unknown. The average degree for the four channels ranges from 1.1 to 2.6.

Using estimated AP locations: We now consider the effect of using estimated AP locations in computing AP degrees. Figure 13 shows the AP interference using four sets of AP locations: actual location of 334 APs, estimated location of those APs detected by war driving, actual location of 521 APs, and estimated location of those APs detected by driving-walking. Not surprisingly, smaller sets (334 APs) have smaller degrees. It is interesting to note that the degree using actual locations is smaller than that using estimated AP locations. This again is because the location of APs are often estimated to be on the path of war driving or war walking, incorrectly placing them closer to one another. We also computed the median of the differences: $median(E_i - A_i)$ where E_i and A_i are the degrees using estimated and actual AP locations, respectively, for AP$_i$. The median for war driving was 2 and the median for driving-walking was 1.

4.7 Summary

We analyzed the effect of using AP locations estimated by war driving and war walking compared to using actual locations. We present a summary of our findings and necessary cautions.

5 Conclusion

The original purpose for war driving was to discover Wi-Fi hotspots. As researchers have started using war-driving traces for other purposes, it is important to better understand errors in war-driving data and the effect of these errors on applications and network characterizations. We collected war-driving traces on the Dartmouth college campus, estimated AP locations from these traces, and compared the estimated locations against actual AP locations. We also analyzed the impact of using estimated locations rather than actual locations on user-location tracking and AP-deployment characterizations. We found that using accurate AP locations is critical in accurately estimating user positions. We observed that estimated AP locations are often biased towards the war-driving paths, which makes the maximum signal range of APs to appear shorter and the interference among APs to appear more severe than in reality. We also found a danger in making assumptions in analyzing traces; even with a conservative assumption that an AP's interference range is 50 m, we still overestimated interference by 12 times. We hope that our study provides necessary cautions in using AP locations estimated by war driving and helps researchers to take necessary steps to cope with errors in the estimates.

Acknowledgments. We would like to thank Nick DeFrancis for war walking on the Dartmouth campus. This project was supported by Cisco Systems, NSF Award EIA-9802068, and Dartmouth's Center for Mobile Computing.

References

1. A. Akella, G. Judd, S. Seshan, and P. Steenkiste. Self-management in chaotic wireless deployments. In *Proceedings of the 11th Annual International Conference on Mobile Computing and Networking (MobiCom)*, pages 185–199, New York, NY, Aug. 2005. ACM Press.
2. N. Bulusu, J. Heidemann, and D. Estrin. GPS-less low cost outdoor localization for very small devices. *IEEE Personal Communications Magazine*, 7(5):28–34, Oct. 2000.
3. Y.-C. Cheng, Y. Chawathe, A. LaMarca, and J. Krumm. Accuracy characterization for metropolitan-scale Wi-Fi localization. In *Proceedings of the Third International Conference on Mobile Systems, Applications, and Services (MobiSys)*, Seattle, WA, June 2005.

4. I. A. Getting. The global positioning system. *IEEE Spectrum*, 30(12):36–38,43–47, Dec. 1993.
5. A. Harter and A. Hopper. A new location technique for the active office. *IEEE Personal Communications*, 4(5):42–47, Oct. 1997.
6. J. Hightower and G. Borriello. Particle filters for location estimation in ubiquitous computing: A case study. In *Proceedings of International Conference on Ubiquitous Computing (UbiComp)*, pages 88–106, Sept. 2004.
7. Kismet. http://www.kismetwireless.net.
8. A. LaMarca, Y. Chawathe, S. Consolvo, J. Hightower, I. Smith, J. Scott, T. Sohn, J. Howard, J. Hughes, F. Potter, J. Tabert, P. Powledge, G. Borriello, and B. Schilit. Place Lab: Device positioning using radio beacons in the wild. In *Proceedings of the 3rd International Conference on Pervasive Computing (Pervasive)*, pages 116–133, Munich, Germany, May 2005.
9. A. LaMarca, J. Hightower, I. Smith, and S. Consolvo. Self-mapping in 802.11 location systems. In *Proceedings of the Seventh International Conference on Ubiquitous Computing (UbiComp)*, pages 87–104, Tokyo, Japan, Sept. 2005.
10. Place Lab. http://www.placelab.org.
11. N. B. Priyantha, A. Chakraborty, and H. Balakrishnan. The Cricket location-support system. In *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking (MobiCom)*, pages 32–43, Aug. 2000.
12. H. Satoh, S. Ito, and N. Kawaguchi. Position estimation of wireless access point using directional antennas. In *Proceedings of International Workshop on Location- and Context-Awareness (LoCA)*, Oberpfaffenhofen, Germany, May 2005. Springer-Verlag.
13. Skyhook Wireless. http://www.skyhookwireless.com.
14. R. Wasinger, A. Krüger, and O. Jacobs. Integrating intra and extra gestures into a mobile and multimodal shopping assistant. In *Proceedings of the 3rd International Conference on Pervasive Computing (Pervasive)*, pages 297–314, Munich, Germany, May 2005.
15. WiFiMaps.com - Wardriving Maps and Hotspot Locator. http://www.wifimaps.com.