A Toolkit for Automatically Constructing Outdoor Radio Maps

Kay Connelly, Yong Liu, Dan Bulwinkle, Adam Miller and Ian Bobbitt
Computer Science Department
Indiana University
{connelly, yonliu, dbulwink, akmiller, ibobbitt}@cs.indiana.edu

Abstract

Outdoor location systems based on Wireless Access Point (WAP) signal strength must either know the exact location of the WAPs in order to use a triangulation algorithm, or must construct a radio-map of the signal strengths. While the radio-map technique increases accuracy and can accommodate a wireless network that is not owned by any one entity, conventional methods for constructing a radio-map are labor-intensive and impractical in such environments. We introduce a toolkit for automatically constructing outdoor radio-maps. Our toolkit can easily be carried by delivery personnel and security guards during their normal work duties to obtain signal readings. The scattered data readings are then fed into an interpolation algorithm to construct a more complete grid that can be used as the radio-map.

Keywords

Wi-Fi, Location System, Radio Map, WAP, GPS

1. Introduction

Location technologies have been widely examined in recent years. The location of a device or person enables location-aware applications such as obtaining directions, finding the nearest resource (be it a restaurant or a printer), and playing physical/virtual games [3]. Indeed, location is an essential input into many ubiquitous computing applications of the future. [6]

Outdoor location systems tend to rely on GPS (Global Positioning System). GPS uses satellite signals to determine location; thus, it does not work indoors and can be affected by such structures as tall buildings in urban areas.

Outdoor location systems using wireless technology such as 802.11 (also known as Wi-Fi) have garnered recent attention for multiple reasons. First, there has been an explosion of wireless-enabled mobile devices. In response to high demand, many PDAs, smartphones and laptops come with built in wireless. GPS, however, rarely comes built-in to such general-purpose mobile devices.

Second, GPS does not work indoors. If one desires a location system both indoors *and* outdoors, GPS is not a complete solution. Given enough wireless access points (WAPs), Wi-Fi location systems can work both inside and outside.

Finally, there has been a dramatic increase in the number of deployed WAPs, resulting in wireless coverage that is dense enough to support outdoor Wi-Fi location systems in many areas. For all of these reasons, Wi-Fi location systems are an attractive alternative to GPS in order to maximize the potential user base and integrate with indoor location technologies.

The major problem with outdoor Wi-Fi location systems, however, is that they either have poor precision or a high overhead in building a radio-map. In this paper, we introduce a toolkit for automatically building a radio-map. The essence of our approach is to combine a GPS unit with a wireless card in a few select devices which are carried around by people who frequently walk a specific area (such as a mailman or security guard). This unit will automatically build a radio-map without knowledge of where the WAPs are located, and can update the radio-map even when WAPs are shut down, added or moved.

In section 2, we discuss the current outdoor Wi-Fi location systems and their limitations. In section 3 we introduce our approach. We provide the technical details of our toolkit in section 4 and our technique for constructing the radio-map in section 5, our experimental results in section 6 and conclude in section 7.

2. Wi-Fi Location Systems

Techniques for locating mobile devices using wireless signal strength have been widely examined for indoor environments [1, 2, 4]. Location systems using such techniques can normally achieve a location accuracy up to 3 meters indoors. Some commercial systems (e.g. Ekahau [5]) can even get a 1 meter accuracy.

Research into Wi-Fi location systems in open outdoor environments has only recently become of interest because of the necessary density of WAPs to make the approach feasible. Within the last two years, many urban and academic settings have added WAPs at an enormous pace. As of October, 2004, downtown Seattle had at least 3,162 access points and the Indiana University Bloomington campus had more than 400 wireless access points within a few square miles.

Wi-Fi location techniques can be divided into two categories: 1. model-based techniques and 2. radio-map techniques.

Model-based techniques make use of the locations of the detected access points and the radio frequency propagation model to triangulate the receiver's position. Model-based techniques are generally considered to have an advantage in outdoor environments since building the radio-map in such environments is a time consuming process. The major disadvantage of this type of location techniques is the lower position accuracy compared to the radio-map techniques.

For example, Place Lab [8, 9] uses the strength of the wireless signals received at a device to locate the "Nearest Access Point" and look up the associated location of that WAP. This provides very coarse-grained location information with an accuracy between 20 and 30 meters.

Rover, on the other hand, feeds the location of the WAPs into a modeling algorithm such as Minimum Triangulation and Curve Fitting [10]. This allows a location to be determined that uses multiple signal strengths. Unfortunately, the paper does not give any measurements of accuracy.

The problem with the model-based approach is that it requires the exact location of all access points. Given the separate ownership and administration of most WAPs, obtaining this information is not trivial. Additionally, without further infrastructure for updating the databases (which are often located on the mobile devices), this approach does not respond well to the dynamic nature of the wireless network with WAPs being added, moved and removed at any time.

The radio-map approach constructs a radio-map by measuring the signal strength of WAPs at multiple points. Radio-map techniques typically require two phases of working. In the offline phase, the signal strengths received from the access points at a set of pre-defined mesh points are gathered and recorded using a database. In the online phase, the currently received signal strengths are compared to the radio-map and the "best" match is returned as the estimated user location. Depending on the density of the network, radio-map techniques can obtain an accuracy of as little as three meters indoors [1].

While this approach does not require the location of the WAPs, current techniques for taking the data points are labor-intensive. Current radio-map approaches also suffer under a dynamic network, requiring the entire map to be re-constructed when WAPs move, join or leave the area.

3. Our Approach

We propose a toolkit that can automatically construct an outdoor radio-map. By equipping a relatively small-number of people with our toolkit, an outdoor radio-map can be updated daily, thereby handling the dynamism encountered by today's wireless networks. Ideal candidates for carrying our toolkit are nomadic in their daily activities (i.e. delivery personnel, such as mailmen and college students who have classes in many buildings).

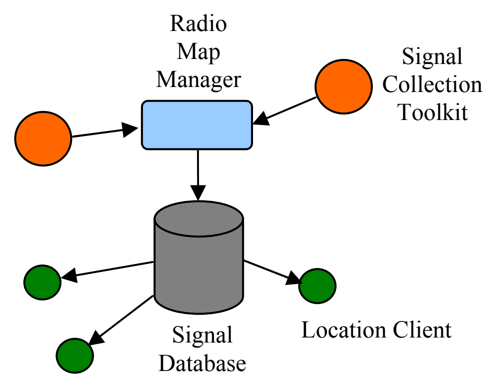

Figure 1. Wi-Fi location platform architecture

Figure 1 shows our wide area Wi-Fi location platform. It consists of four major components:

1. Signal collection toolkit
2. Radio-map manager
3. Signal database
4. Location client

The signal collection toolkit equips a PDA with both GPS and a wireless access card. As a person carries the toolkit, it records the wireless signal strengths along with the GPS coordinates automatically. The person carrying the toolkit does not need to interact with the system in any way. They act as data collectors for the wide area Wi-Fi location system in the offline phase.

The frequency of the signal readings can be tailored. The accuracy of our constructed radio-map is primarily limited by the resolution of the GPS module, which is 1 meter. This allows our radio-map approach to have a much higher resolution than existing model-based Wi-Fi location approach.

The radio map manager is in charge of building the comprehensive radio map for a specific area from a set of signal strength readings and their associated locations. The radio-map manager feeds the newly generated radio map to the signal database. For every location reading, the signal database holds a set of signal strengths: one signal strength for each WAP that can be detected at that location.

In order to determine their location, devices equipped with a wireless access card may use the location client. The location client reads all available signal strengths and submits the set to the database, which can then retrieve a location based on the nearest point in the radio-map.

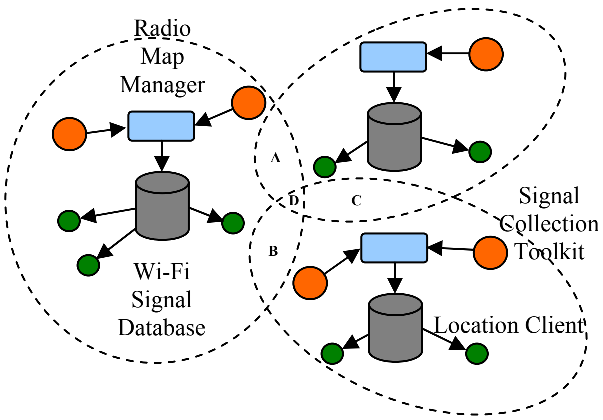

Figure 2. Scalability of the wide area Wi-Fi location platform

Figure 2 shows how to scale our approach. Similar to the GSM network in which there are cells of coverage, a large area can be split into multiple cells. Each cell would have its own radio-map managers and signal databases. When a mobile device comes into a place where the coverage of two or more signal databases overlaps (such as area A, B, C, and D in Figure 2), the location client may connect to any of the databases to retrieve its current location data.

4. Signal Collection Toolkit

Our signal collection toolkit is implemented on an iPAQ 4155 handheld device running Windows Mobile 2003. This toolkit includes two major components:

- Wi-Fi signal strength reading collection
- GPS reading collection

In the Wi-Fi signal strength reading collection component, Microsoft Windows CE's Network Driver Interface Specification (NDIS) is used to implement a low level call to Device IO Control to obtain the MAC address and signal strength for each WAP in range. The GPS unit collects standard NMEA[1] sentences. Our GPS reading collection component parses the $GPGGA sentences to obtain the latitude and longitude.

The data collected by the GPS unit, paired with the wireless signal strengths and MAC address readings, is stored in a file until the PDA is synched with the desktop. Once synched, the data is automatically sent to the radio map manager. The radio-map manager periodically re-builds the comprehensive radio map for the entire area based on the readings inputted by the toolkit and keeps updating the database.

For our current implementation, the frequency of Wi-Fi and GPS readings are set to 1 reading per second. In our implementation, we pair the latest GPS reading with the current Wi-Fi reading. Thus, it is important to note that the current implementation of this toolkit targets pedestrians.

5. Signal Strength Interpolation

Current approaches for constructing radio-maps collect data at specifics points (i.e. a 1-meter grid). While our system greatly reduces the labor involved in data collection, we cannot easily collect data at pre-defined locations because of its automated nature. As we have no control over where the data collectors will walk, we have to be able to handle holes in the data.

To deal with the non-uniform distribution of data points, we utilize a data interpolation method called inverse distance weighting (IDW) [7] to build the comprehensive radio-map. In this way, we use our collected scatter points to interpolate a value for every point on the grid. This interpolated grid of points is then used in a normal radio-map algorithm.

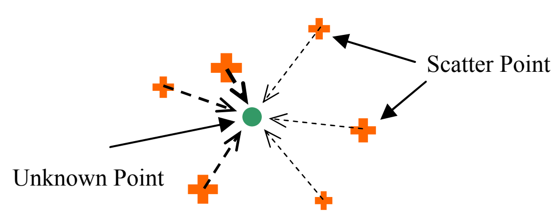

Figure 3. Inverse distance weighting

The inverse distance weighting method is based on the assumption that the interpolation points should be influenced most by their nearest neighbors and less by the more distant points. The interpolating surface is a weighted average of the scatter points and the weight assigned to each scatter point diminishes as the distance from the interpolation point to the scatter

---

[1] National Maritime Electronic Association

point increases. Figure 3 illustrates this method in a conceptual level.

A simple function used to compute the interpolating surface using the inverse distance weighting method is as the following:

$$F(x,y)=\sum_{i=1}^{n} w_i f_i,$$

where n is the size of the scatter point set, f_i is the value corresponding to each scatter point, and w_i is the weight assigned to each scatter point. The weight of each scatter point can be calculated using the following formula:

$$w_i=\frac{h_i^{-p}}{\sum_{j=1}^{n} h_j^{-p}}$$

where p is the power parameter (an arbitrary positive integer, typically p=2) and h_i is the distance from the scatter point to the interpolation point. By definition, the sum of all the weights is 1.

A major limitation of using the IDW method with the scattered data points as the only input is that the interpolating surface is a simple weighted average of the data values of the scatter points and is constrained to lie between the extreme values in the data set. For outdoor Wi-Fi signal strength reading collection, this means the signal strengths of all the interpolated points will be at least equal to the lowest collected signal strength. But since the Wi-Fi signal coverage is typically 100 meters, at any point with a distance to the access point larger than 100 meters, the received signal strength should obviously be close to 0.

Our toolkit, however, only reports detected signal strengths. Thus, we must adjust our computation by introducing artificially generated data points that correspond to no signal being detected. The weakest signal strength our wireless card can detect is -97 dBm[2]. Thus, for every point in the database, if a signal strength is not recorded for a particular WAP, we set it to be -100 dBm.

Further, since a Wi-Fi signal does not travel more than 100 meters, it does not make sense to include a reading which is very distant from the point that is being interpolated. This greatly slows the computation for a point which will have a weight approaching zero. Thus, we define a radius, r, which is used to determine if a measured location will be used to compute a point in the radio map. Only when a mesh point is within the radius of the signal reading will the reading be used in the IDW algorithm for that point. For our experiments, we set r to 50 meters.

Figure 4 shows how we apply the improved IDW algorithm to generate the comprehensive radio-map for a target area. In Figure 4, a radio map grid with 9 by 12 points is generated for a single access point. The access point is denoted by an orange point near (3, I) in the figure. In this target area, nine Wi-Fi signal readings have been gathered for this access point during preliminary signal collection. The locations at which these readings are collected are denoted by orange + symbols in the figure. The larger the symbol is, the stronger the signal detected at this point is. The black **x** symbols denote the locations at which we generate the artificial signal readings of -100 dBm.

In the figure, we have drawn the radius around three points near: (4, F), (6, L) and (7, E). When computing the green point at (6, G), the signal reading near (6, L) does not contribute to its strength computation since this reading is too far away from the green point.

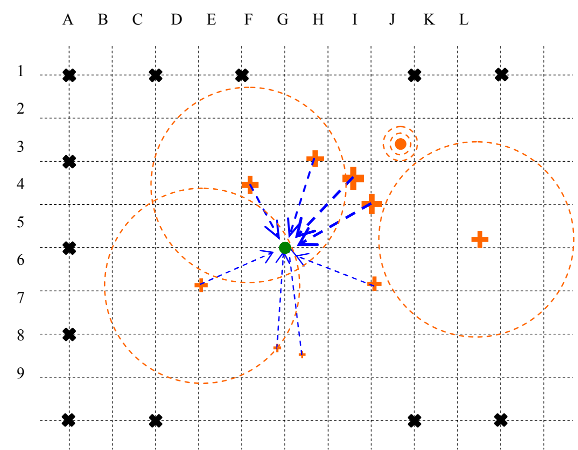

Figure 4. Illustration of improved IDW algorithm

In the wide area Wi-Fi location platform, the signal strength interpolation for a specific area is conducted periodically. When new access points are detected, they will be automatically folded in during the interpolation process. In addition, an aging algorithm can be introduced to help purge those old access points which are not active any more. The Wi-Fi signal database is then updated readily with little effort.

6. Experiment Results

In our experiments, we collect 13,657 raw Wi-Fi signal and GPS reading pairs on the Indiana University Bloomington campus using the Wi-Fi signal collection toolkit. All of the data was collected within a 10,000

---

[2] The strongest signal strength we recorded was -55 dBm. Signal strength about 1 meter away from the WAP can be as good as -20 dBm.

square-meter area on campus in which 11 academic and student activity buildings are located. An aerial photo of this area is shown in Figure 5.

The total time spent in data collection was two hours. During this time, a researcher carried the signal collection toolkit and walked along the paths among the buildings several times. There was a total of 78 wireless access points detected. Signals from these access points covered most of the area, except for a few places in the woods surrounded by the buildings. In Figure 5, the positions of the readings are denoted with different colored and shaped symbols indicating different WAPs.

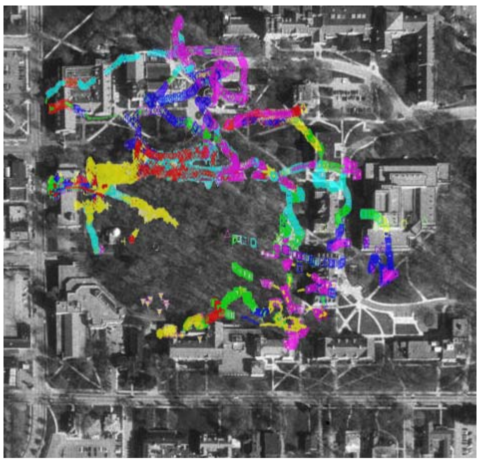

Figure 5. An aerial photo marked with Wi-Fi signal readings

After the data collection was complete, we used the improved IDW algorithm to generate the interpolation points from the collected signal readings. It took 134 minutes to generate a 1.1 meter grid.

Figure 6 shows a sample signal strength interpolation for a single access point. The x-axis denotes longitude and the y-axis denotes latitude. The z-axis denotes the strength of the received Wi-Fi signal with dBm as the unit. The red points in the figure are the initial scatter points which were collected using the signal collection toolkit. The black points on the x-y plane are the artificially generated scatter points, which help to improve the interpolation results. In this sample interpolation process, the granularity of the artificially generated points was 10 meters. From experiments, we found a finer granularity of these fake points did not improve the interpolation results. On the contrary, it significantly slowed down the interpolation process.

Based on the initial scatter points and the artificially generated points, a fine-grained interpolation surface is generated by applying the improved IDW algorithm. The granularity of the surface in our experiments is 1 meter.

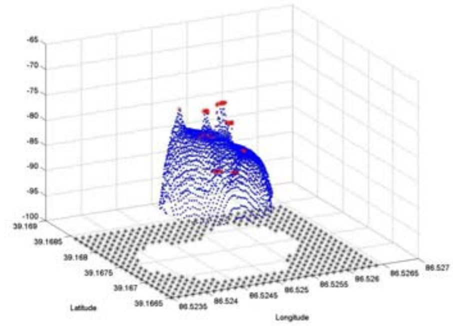

Figure 6. Interpolation points for a single WAP

From Figure 6, we can see an interpolation surface is generated only for a limited area near the initial scatter points. This is because the signal strength at the interpolated points outside this area falls below a threshold, which is -90 dBm in our experiments, and hence they are deleted from the database.

We performed a cross-validation study to examine the interpolation accuracy from the initial scatter points. We divided the entire set of signal strength readings for a specific access point into 100 subsets. During each iteration of the cross validation, we use one subset as the test set and the remaining 99 subsets as training sets. After computing the interpolated points with the training sets, we examined the locations of our test set and compared the actual signal strengths to the interpolated values. We repeated this process 100 times and computed the error rate for each iteration.

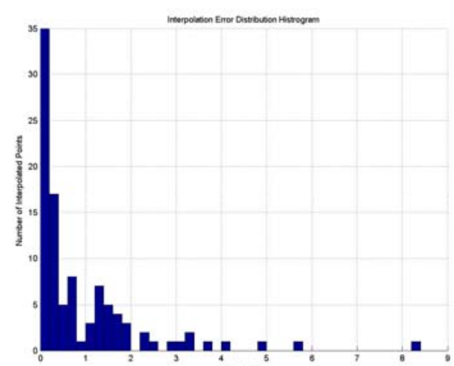

Figure 7. Interpolation error distribution

The average interpolation error is 0.98 dB. Since the outdoor Wi-Fi signal strength detected in our experiment ranges from -55 dBm to -97 dBm, the relative interpolation error is only 2%. Figure 7 shows the error distribution of our sample signal strength interpolation. The x-axis is the error in dBm, and the y-axis is the number of interpolated points. From this figure we see that the majority (52%) of the interpolation errors are within 0.4 dB.

When determining a location, multiple WAPs are typically used, which can offset the few points with poor interpolation results. For our experiment, when the grid resolution is set to 1 meter, we had a total of 3700 points with Wi-Fi readings and 331 points with no WAP within range. On average, 2.65 WAPs were detected per reading, with as many as 9 WAPs.

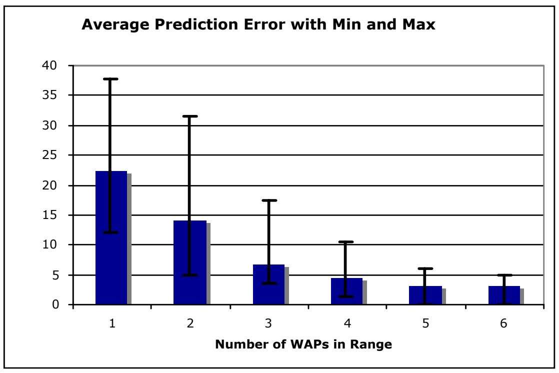

Figure 8. Error based on # of WAPs

We performed a test on the accuracy of our generated radio map based on the number of WAPs in range. Figure 8 gives the average, minimum and maximum error for 1-6 detected WAPs, using 10 locations each. As the figure shows, detecting at least 4 WAPs provides an average error of less than 5 meters and a maximum error of 10.4 meters. Detecting 2-3 WAPs gives an error less than the Place Lab algorithm on average. Further, the *worst* error computed with 3 WAPs is also better than Place Lab. At locations where only a single WAP is detected, the location accuracy is similar to the accuracy of a "Nearest Access Point" location system.

The test result clearly shows that the higher the density of WAPs in an area, the smaller the error will be. Thus, the number of detected WAPs could be used to provide an application or end-user with a confidence value of the predicted location.

7. Conclusion

Radio-map approaches for outdoor location system based on wireless signal strength have not previously been pursued because of the cost in constructing the map. In this paper, we introduce a toolkit for automatically collecting the signal strengths, without any human intervention. While our toolkit is labor-free, it can only gather scattered data points. We introduce an interpolation technique that can handle scattered points to build a comprehensive radio-map. We performed cross-validation on our data set and show an average error of less than 1 dB.

Further, we analyzed the accuracy of our radio-map predictions based on the number of WAPs in range, and showed that the higher the density of WAPs, the smaller the error. Our results can be used to inform the end user or location-based application of the confidence of the current location prediction.

In the next phase, we plan on deploying a working wide area Wi-Fi location system with a set of location-aware applications supported on Indiana University Bloomington campus. We believe this wide area location system will greatly enhance the user experience for various location-aware applications.

References

[1] P. Bahl and V. N. Padmanabham. RADAR: An In-Building RF-Based User Location and Tracking System. In *Proceedings of Infocom 2000*, pp. 775-784, March 2000.

[2] P. Castro, P. Chiu, T. Kremenek, and R. R. Muntz. A probabilistic room location service for wireless networked environments. In *Proceedings of the Third International Conference on Ubiquitous Compute (Ubicomp)*, Atlanta, GA, pp. 18-34. 2001.

[3] M. Chang and E. Goodman. FIASCO: game interface for location-based play, In *Proceedings of ACM Designing Interactive Systems*, pp. 329-332, 2004.

[4] R. Doshi, T. Jagatic, A. Khalil, S. Shirasuna, and P. Venkatakrishnan. Development of a Wireless Location System in Lindley Hall. Technical Report #589, Department of Computer Science, Indiana University, Bloomington, IN, 2004.

[5] Ekahau Positioning Engine. http://www.ekahau.com

[6] J. Hightower and G. Borriello. Location systems for ubiquitous computing. *IEEE Computer*, pp. 57-66, August 2001.

[7] Inverse Distance Weighting. http://www.ncgia.ucsb.edu/pubs/spherekit/inverse.html

[8] Place Lab. http://www.placelab.org.

[9] B. Schilit, et al. Challenge: Ubiquitous Location-Aware Computing and the Place Lab Initiative. In *Proceedings of The First ACM International Workshop on Wireless Mobile Applications and Services on WLAN (WMASH 2003),* San Diego, CA. pp. 29-35, September 2003.

[10] A. U. Shankar, et al. Rover technology: Enabling Scalable Location-aware Computing. *IEEE Computer*, pp.46-53, 2002.