Location Accuracy Estimates for Signal Fingerprinting

John Krumm
jckrumm@microsoft.com
Microsoft Research
Redmond, WA

ABSTRACT

Location fingerprinting is a technique for determining the location of a device by measuring ambient signals such as radio signal strength, temperature, or any signal that varies with location. The accuracy of the technique is compromised by signal noise, quantization, and limited calibration resources. We develop generic, probabilistic models of location fingerprinting to find accuracy estimates. In one case, we look at predeployment modeling to predict accuracy before any signals have been measured using a new concept of noisy reverse geocoding. In another case, we model a previously deployed system to predict its accuracy. The models allow us to explore the accuracy implications of signal noise, calibration effort, and quantization of signals and space.

CCS CONCEPTS

• **Information systems → Location based services**; **Geographic information systems**; **Mobile information processing systems**.

KEYWORDS

location, signal fingerprinting, geocoding, noisy reverse geocoding, indoor location, WiFi location

ACM Reference Format:
John Krumm. 2020. Location Accuracy Estimates for Signal Fingerprinting. In *28th International Conference on Advances in Geographic Information Systems (SIGSPATIAL '20), November 3–6, 2020, Seattle, WA, USA.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3397536.3422243

1 INTRODUCTION

Fingerprinting is a common technique for determining location. By measuring ambient signals such as radio signal strength or temperature, a device can localize itself down to some degree of spatial resolution. Fingerprinting is an alternative to GPS that can work indoors and without the startup time and power costs of GPS.

Accuracy models of fingerprint location systems are important for understanding their limits and for configuring them to meet specifications. Anticipating accuracy can inform decisions on whether or not to deploy as well as help estimate what levels of effort and expense will be required to build the system at the desired level of performance.

SIGSPATIAL '20, November 3–6, 2020, Seattle, WA, USA

ACM ISBN 978-1-4503-8019-5/20/11…$15.00
https://doi.org/10.1145/3397536.3422243

Location-based fingerprinting associates measured, ambient signals with spatial regions. To increase accuracy and coverage, there may be multiple measured signals, such as from multiple radio transmitters. The process converts noisy input signals into uncertain location inferences. This paper models the conversion process, and in particular maps the noise level of the input to the location uncertainty of the output. Besides signal noise, the models also account for signal quantization, spatial quantization, the number of signals, and the calibration effort. We demonstrate how to use the models to anticipate how location accuracy varies with the parameters of the system.

2 PROBLEM SETTING AND RELATED WORK

Location fingerprinting translates a measured signal into a location. The signal may be one-dimensional or more. As an example, we can imagine a mapping between outside temperature and location. Given a measured temperature, the set of possible locations on the earth is reduced. We could add more measurements, such as the strength of gravity [7], outside brightness, and the angle of the sun in the sky [4] to make a multidimensional signal that would further reduce the set of possible locations. Another common example is the signal strength of localized radios, such as WiFi [1], cell towers [14], or commercial FM [9].

This paper develops two accuracy models. The first, in Section 3, is a predeployment model. It is intended to model fingerprinting accuracy before any signal generating devices, such as WiFi access points, are deployed, which is convenient for planning a deployment and understanding accuracy limits. The predeployment model divides space into distinct regions naturally based on the full range of the available discrete signals. We introduce the concept of noisy reverse geocoding for predeployment. The second model, for post-deployment in Section 5, is aimed at fingerprint-based location systems where the spatial regions have already been defined both geometrically and by their representative signal strength vectors. Both models are applicable to any type of signal that varies with location.

While there has been much work on fingerprint-based location, especially for WiFi [2], there has been comparably little work on theoretical analysis of such systems. One exception is the work of Kaemarungsi and Krishnamurthy [8]. Their work focuses on WiFi, and they model signal noise as Gaussian, like us. They develop a probabilistic model of signal matching on a predetermined grid, testing with simulated signal strengths. Our predeployment model differs from theirs in that we account for inevitable signal quantization as well as the effects of multiple readings during the calibration phase. Our predeployment model also works independently of a particular discretization of space in an attempt to generalize performance limits.

Another related model, also aimed at WiFi, is that of Battiti, Brunato, and Delai [3]. Their goal was to place WiFi access points such that network coverage and location accuracy were both optimized. Like us, they account for signal noise and quantization, but their model assumes a predetermined grid of test points, while our predeployment model generalizes away this assumption. This allows us to reason about location accuracy vs. spatial resolution.

Both [8] and [3] are tested on simulated signal strengths. In contrast, our postdeployment model is tested on actual data.

Our accuracy models are designed to be general, not relying on propagation models nor spatial continuity assumptions about the fingerprinted signals. For instance, applied to WiFi, our approach does not depend on a path loss model. This means our performance predictions are independent of the particular spatial variations of the measured signals. Without this assumption, we model space as a collection of discrete regions, each represented by a distinct signal strength vector. Because of this, our performance metric is classification accuracy, which is the probability that a set of signals measured during runtime will be associated with the correct discrete region. This is the same performance metric used by [8] and [3] in their WiFi fingerprinting models.

3 NOISY REVERSE GEOCODING

We first build a predeployment model, covering cases where no signals have been measured. We develop a postdeployment model in Section 5. The predeployment model is a variant of geocoding, which is the act of assigning unique identifiers to the discrete spatial cells covering a region. One example is postal codes which identify irregular polygons on the earth's surface with strings of characters. A modern example is what3words that assigns a unique three-word string to each 3m × 3m cell on the surface of the earth [15]. For instance, the Space Needle in Seattle, WA is geocoded with the three words "clear.sheets.title". A geocode can be thought of as an n-digit identifier with an alphabet of size n_j for each digit. In the case of what3words, $n = 3$ and $n_1 = n_2 = n_3 \approx 40,000$, where $40,000$ is the number of possible words for each word of the geocode. The number of possible geocodes is $N = \prod_{j=1}^{n} n_j$. For what3words, $N = 40,000^3 = 6.4 \times 10^{13}$, which is close to the 5.7×10^{13} 3m × 3m cells that what3words uses to tile the earth.

Reverse geocoding is the translation of a geocode into an actual location. For instance, a list of WiFi base stations and their signal strengths could be translated into a region on earth where we expect to measure those signal strengths from those base stations. This is a convenient way to measure the location of a mobile device [1]. The alphabet for each base station is the list of possible discrete signal strengths, normally measured in integer dBm. Noisy reverse geocoding occurs when the measurement is noisy, leading to uncertainty in the geocode. We are interested in translating this geocode uncertainty into spatial uncertainty. Our model accounts for intrinsic signal noise and measurement noise, signal quantization effects, spatial quantization, and calibration effort.

3.1 Measurement Distributions

Mathematically, we will designate a geocode representing a spatial region $\mathcal{R}_i$ as a vector of length n denoted by boldface $\mathbf{m}_i^{(r)}$. (The (r) superscript stands for "runtime", which we explain below.) The elements of $\mathbf{m}_i^{(r)}$ are the scalars $m_{i,j}^{(r)}$, $i \in 1...N$ and $j \in 1...n$. To reiterate, i indexes the N physical regions on the ground, and j indexes the n scalar elements of the vector $\mathbf{m}_i^{(r)}$. The number of possible values for $m_{i,j}^{(r)}$ is the size of the alphabet n_j. For WiFi, the alphabet is the list of possible discrete signal strength values. For temperature, the alphabet is the list possible discrete temperature values.

The discrete geocode vector comes from discretizing a continuous signal, i.e. the discrete $m_{i,j}^{(r)}$ comes from the continuous $s_{i,j}$. We will assume that the intrinsic noise of the signal, combined with the measurement noise of the mobile device, is represented by a Gaussian distribution such that $s_{i,j} \sim \mathcal{N}\left(\mu_{i,j}, \left(\sigma_{i,j}^{(r)}\right)^2\right)$. The Gaussian distribution function is given by

$$g(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

Then the discrete measurement $m_{i,j}^{(r)}$ is produced by a uniform quantizer such that its probability is

$$P^{(r)}\left(m_{i,j}^{(r)}\right) = \int_{m_{i,j}^{(r)}-\frac{1}{2}}^{m_{i,j}^{(r)}+\frac{1}{2}} g\left(s; \mu_{i,j}, \left(\sigma_{i,j}^{(r)}\right)^2\right) ds \tag{1}$$

This says that region $\mathcal{R}_i$ may produce different discrete measurements $m_{i,j}^{(r)}$, due to the noise in $s_{i,j}$. This is the same quantization model as used in Battiti et al. [3] for WiFi signal strengths.

Fingerprinting refers to the act of gathering signals inside spatial regions $\mathcal{R}_i$ whose locations are known. The fingerprints serve as a reference map for reverse geocoding runtime measurement vectors $\mathbf{m}^{(r)}$ into inferences about the measurement device's current region. While these fingerprints can come from simulation (e.g. radio propagation for WiFi location), it is usually more reliable to gather fingerprints manually in a calibration phase, such as the original RADAR system for WiFi location [1]. In Section 4.2 we show how a simple radio propagation model for WiFi is a poor predictor of signal strength. A typical scenario is to use sensors to measure the signals. We refer to these fingerprints as $\mathbf{m}^{(c)}$, where the superscript "c" stands for "calibration". They are also uniformly quantized similarly to the runtime signals in Equation 1:

$$P^{(c)}\left(m_{i,j}^{(c)}\right) = \int_{m_{i,j}^{(c)}-\frac{1}{2}}^{m_{i,j}^{(c)}+\frac{1}{2}} g\left(s; \mu_{i,j}, \left(\sigma_{i,j}^{(c)}\right)^2\right) ds \tag{2}$$

The major difference between quantizing for calibration (Equation 2) and runtime (Equation 1) is the signal noise variance $\left(\sigma_{i,j}^{(c)}\right)^2$ vs $\left(\sigma_{i,j}^{(r)}\right)^2$. The calibration variance is likely smaller, due to the extra care taken to make it accurate. One typical scheme is for the calibration technician to take the mean of l signal measurements, while the runtime signals come from just a single measurement. For Gaussian distributed signals, this means that $\sigma_{i,j}^{(c)} = \sigma_{i,j}^{(r)}/\sqrt{l}$.

We assume that the n constituent scalar signals in the vectors are statistically independent, which means the probabilities of the discrete signal vectors are given by products of the probabilities of the discrete scalar probabilities:

$$P^{(r)}\left(\mathbf{m}_i^{(r)}\right) = \prod_{j=1}^{n} P^{(r)}\left(m_{i,j}^{(r)}\right)$$

and

$$P^{(c)}\left(\mathbf{m}_i^{(c)}\right) = \prod_{j=1}^{n} P^{(c)}\left(m_{i,j}^{(c)}\right)$$

This independence assumption is justified by the normally independent nature of the constituent signals.

We further assume statistical independence between the calibration and runtime phases. Thus the joint probability distribution of the calibration and runtime measurements is

$$P^{(c),(r)}\left(\mathbf{m}_i^{(c)}, \mathbf{m}_i^{(r)}\right) = P^{(c)}\left(\mathbf{m}_i^{(c)}\right) P^{(r)}\left(\mathbf{m}_i^{(r)}\right) \tag{3}$$

3.2 Probability of Correct Classification

Equation 3 gives the joint probability distribution of the calibrated and runtime signal strengths for a spatial region. The runtime measurement will be properly assigned to the correct region if its measurement vector $\mathbf{m}^{(r)}$ matches the region's calibration vector $\mathbf{m}^{(c)}$. Thus the probability of a correct match is

$$P\left(\mathbf{m}_i^{(r)} = \mathbf{m}_i^{(c)}\right) = \sum_{i=-\infty}^{\infty} P^{(c),(r)}(\mathbf{m_i}, \mathbf{m_i}) \tag{4}$$

We will use the probability of correct classification as the accuracy metric in this paper. This is the same performance criterion used by Battiti et al. [3] and Kaemarungsi and Krishnamurthy [8] in their models of WiFi fingerprinting performance.

For multidimensional signals, such as WiFi signal strengths, there will be many possible measurement vector values, specifically N possible vectors, one from each spatial region $\mathcal{R}_i$. Unless the calibration and runtime measurements are very precise, the multitude of small spatial regions would lead to low classification accuracies. For this reason, it is attractive to combine regions so each is represented by multiple fingerprints $\{\mathbf{m}_{i+\delta}^{(c)}\} : \delta \in D$, where D represents index offsets to the calibration vectors of nearby spatial regions. For the sake of exposition, we will group regions using a set of indices $D(\Delta, n)$ which represents a total of $(2\Delta+1)^n$ index offsets representing a hypercube of nearby regions. (As a reminder, n is the signal dimensionality.) In reality, a technician would likely group together regions that are physically or semantically close to each other. As a simple grouping scheme, we can rewrite the probability that a measurement would be part of the calibration signals that make up the group:

$$P\left(\mathbf{m}_i^{(r)} \in \{\mathbf{m}_{i+\delta}^{(c)}\} : \delta \in D(\Delta, n)\right) = \sum_{i=-\infty}^{\infty} \sum_{\delta \in D(\Delta,n)} P^{(c),(r)}(\mathbf{m}_{i+\delta}, \mathbf{m_i}) \tag{5}$$

This is simply combining spatial regions into groups of $(2\Delta+1)^n$ spatial regions. This means the total number of regions becomes $N_\Delta = N/(2\Delta+1)^n$.

3.3 Air Temperature Example

Here we present a simple example of noisy reverse geocoding using ambient outdoor temperature. A mobile thermometer could measure outdoor temperature and then consult a real-time table to determine which regions on the earth currently have that temperature. In this case, the signal has only one dimension, so $n = 1$. The earth's minimum and maximum recorded temperatures are -89.2° C[1] and 56.7° C[2] [12]. Rounding to integers, the alphabet of temperatures is [-89, 57], meaning $N = n_1 = 147$. With the earth's entire surface area of 510×10^6 km^2, the average size of each region would be about 3.47×10^6 km^2, which is roughly the size of India. These regions would not necessarily be connected. Clearly temperature is not a high-precision location signal, but it serves as a simple example, and it could be combined with other signals.

Thermistors are commonly used as thermometers in electronic devices, and one study found their standard deviations in temperature measurement varied from 0.067° C to 0.12° C [11]. For the sake of the example, we will say the calibration is done by existing ground stations with the higher accuracy value, i.e. $\sigma_{i,j}^{(c)} = 0.067°C$, and the runtime measurements on mobile devices are done with the lower accuracy value, i.e. $\sigma_{i,j}^{(r)} = 0.12°C$. From Equation 4, this gives a classification accuracy of 0.999969. It is possible that the fingerprinted temperatures may be interpolated or predicted via weather forecasting, leading to a lower calibration precision $\sigma_{i,j}^{(c)}$. Figure 1 shows that classification accuracy drops as calibration precision becomes worse, as expected. While it may seem that a temperature-based location system is not very flexible, given that we cannot control outdoor temperatures, this analysis shows that we can still control the system's accuracy by making better temperature measurements.

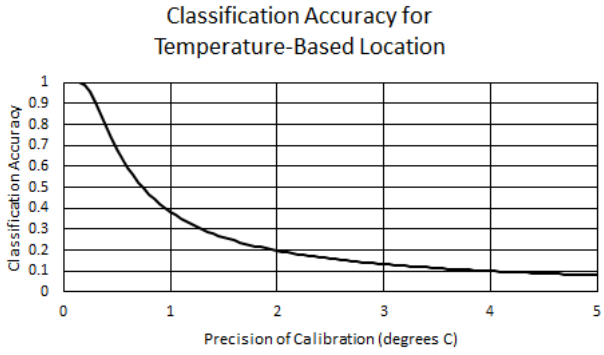

Figure 1: As fingerprinting calibration precision $\sigma_{i,j}^{(c)}$ becomes worse, the classification accuracy drops according to our mathematical model.

3.4 WiFi Location Example

For this example, we imagine $n = 3$ WiFi base stations at the three corners of an equilateral triangle, as shown in Figure 2. Each base

[1] Vostok, Antarctica, July 21, 1983
[2] Furnace Creek, CA, USA, July 10, 1913

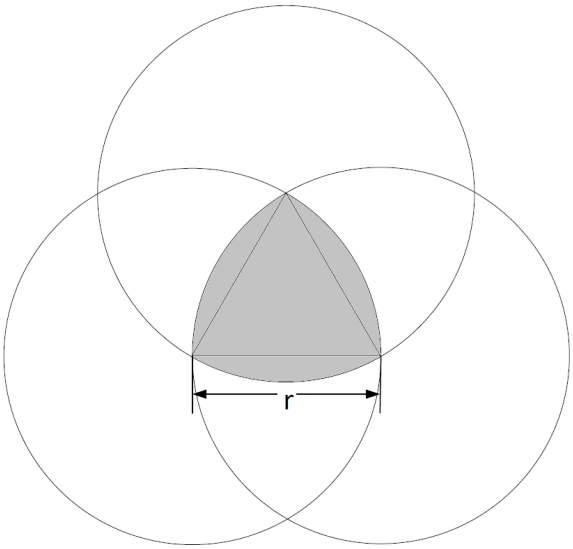

Figure 2: The Reuleaux triangle, in gray, represents the region of three overlapping WiFi ranges, where each range is a circle of radius r.

station has a range of r, and we assume each side of the triangle has length r. The region where all three base stations are detectable is a Reuleaux triangle, which is shown in gray in Figure 2. The area of the overlap is $A = 0.5(\pi - \sqrt{3})r^2$, and we assume there is a mobile device inside this region whose location we want to compute.

We will use realistic WiFi parameters from [13] to determine values for our model. The authors of [13] measured WiFi signal strengths in a building to assess the feasibility of using WiFi for indoor location. In one experiment shown in their Figure 6, they found that signal strengths varied over [-89,-46] dBm, giving $n_1 = n_2 = n_3 = 44$ and $N = \prod_{j=1}^{n} n_j = 85,184$ distinct geocodes. A reasonable assumption for the range of WiFi is about 200 feet or $r = 60.96$ meters. This gives the Reuleax triangle an area of $A = 2,619\,\text{m}^2$, and an average area covered by each geocode of $A/N = 0.031\,\text{m}^2$. If this small area were a circle, its radius would be about 10 cm.

Using a weak statistical test, the authors of [13] found that the measured WiFi signal strengths at a static location were Gaussian distributed with a standard deviation of 2.13 dBm. The authors of [5] also found a Gaussian distribution for WiFi signal strengths, and [3] and [8] use this assumption as well. Thus we will assume mobile devices measure signal strengths from a Gaussian distribution with a runtime $\sigma_{i,j}^{(r)} = 2.13$ dBm. We will also assume that the space is calibrated by taking $l = 5$ signal strength readings to geocode every region, giving $\sigma_{i,j}^{(c)} = \sigma_{i,j}^{(r)}/\sqrt{l} = 0.95$ dBm. We note that geocoding all 85,184 regions this way would be very tedious if done manually.

From Equation 4, the probability of correct classification is 0.0048. This is small, but still about 407 times more accurate than choosing one of the N regions uniformly at random.

3.4.1 Improving Calibration. One way to improve the classification accuracy may be to calibrate more carefully by averaging together more signal strength measurements for each geocode. The classification accuracy above assumed $l = 5$ calibration measurements for every region. Figure 3 shows what happens when the number of calibration readings is varied. We see that accuracy increases significantly up to about $l = 20$ readings, but only slowly after that. The asymptote is at a classification accuracy of around 0.0065, which is still low, so increasing calibration effort may not be a good tactic.

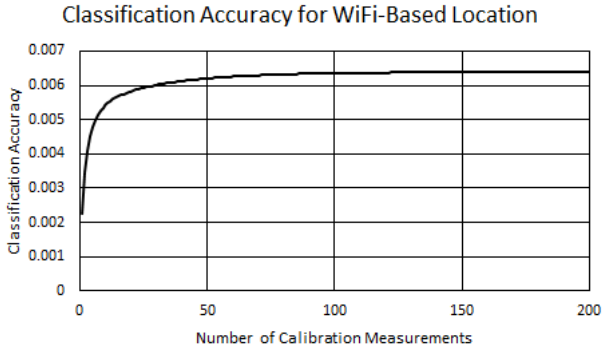

Figure 3: Based on our model, classification accuracy grows when there are more WiFi measurements used for calibrating each region.

3.4.2 Consolidating Regions. The natural region size, as discussed above, is the coverage area divided by the number of distinct, possible signal strength vectors. This size is relatively small in this example. For increased classification accuracy, it may be worthwhile to reduce the granularity of the space by increasing the region size. Each region would be associated with multiple calibration vectors. Equation 5 shows how to compute the new classification accuracy in this case. We do this by consolidating smaller nearby regions into groups of larger regions. The grouping is controlled by Δ, and each group is made up of $(2\Delta + 1)^n$ of the smallest regions, with $n = 3$ in our case of three WiFi base stations.

Figure 4 shows the results for different values of Δ. As Δ grows, the mean size of the regions grow, as shown with the gray curve. For instance, at $\Delta = 4$, the mean region size is $22.4\,m^2$. With a smaller number of regions comes better classification accuracy: at $\Delta = 4$, the classification accuracy is 0.842. The classification accuracy continues to grow with Δ. This shows how our model expresses the tradeoff between region size resolution and classification accuracy, and it is helpful for choosing an operating point that satisfies both criteria.

3.4.3 Adding Another Signal. Our model lets us show the effect of adding another signal for localization. Imagine adding another WiFi base station of the same radius at the center of the Reuleaux triangle in Figure 2. The total coverage area would stay the same, but the number of signals would rise to $n = 4$. Using the same analysis as in Subsection 3.4.2, we can compute the region classification accuracy

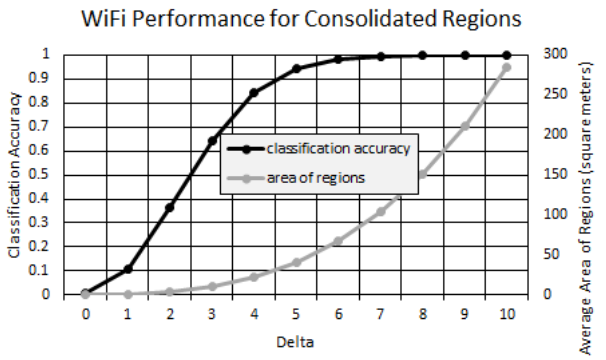

Figure 4: Classification accuracy and average region size both grow as smaller regions are grouped into larger regions, according to our model.

as a function of the size of the regions. As shown in Figure 5, classification accuracy grows noticeably after adding the new base station. For instance, at a spatial resolution of around 25m^2, three base stations give a classification accuracy of about 0.842, while four base stations give about 0.990. At the same resolution, reducing to two base stations would give a classification accuracy of about 0.270. These performance trends match our intuition that adding more signals gives more accuracy.

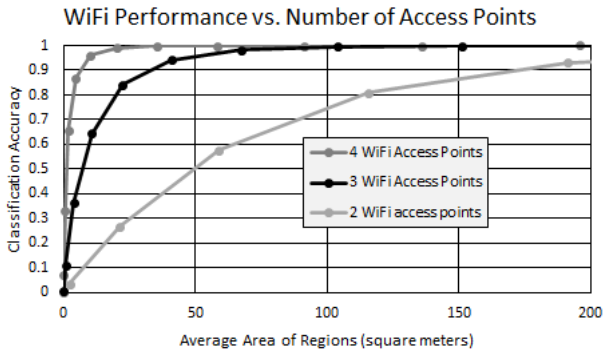

Figure 5: Theoretical classification accuracy improves with the addition of more base stations.

Our predeployment model thus accounts for signal noise, signal quantization, calibration effort, and the number of distinct signals. These two examples show how to use our model to anticipate the effects of different choices on the performance of the system before deployment. Before we discuss our postdeployment model, we explain the experimental data we use to verify the postdeployment model in the next section.

4 EXPERIMENTAL DATA

This section describes the experimental data we used in the next section on Voronoi cell classification.

4.1 WiFi Data

The experimental data comes from Mendoza-Silva et al. [10] who gathered monthly WiFi signal strength data for 25 months, available for free download. A trained professional collected the data from two floors of a library at a set of consistent reference point locations in the building. The reference points for both floors had the same relative coordinates, shown in Figure 6. We used data from the fifth floor where the data collector was facing in the "up" direction, i.e. toward the top of the floor plan. For each collection session, there were six signal strength scans taken in immediate succession from each location and each facing direction. Each scan returned the signal strengths of all WiFi base stations in range. The paper raises a suspicion that the first of each group of six readings was actually buffered from the previous scan location/orientation, so we dropped the first reading in every group of six. To simplify our analysis, we considered only the top three most frequently detected WiFi base stations, which were numbers 16, 50, and 51. We used a total of $N = 72$ different reference points, which are the green, blue, and magenta points in Figure 6. For each reference point, there was an average of about 125 sample signal strength vectors that contained signal strengths from all three WiFi base stations. The number of vectors per reference point varied from 12 to 304, and the number for each reference point is shown in Figure 7. Figure 8 shows a histogram of signal strength values, which varied from -97 to -33 dBm.

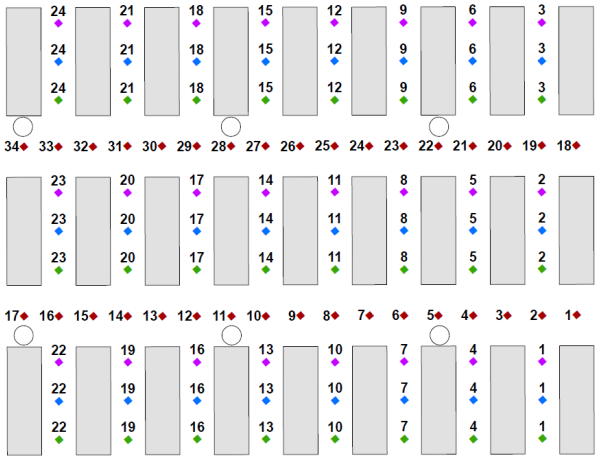

Figure 6: These are the reference points in the experimental dataset from [10]. We used all but the two horizonal lines of points in red. (Figure copied with permission.)

4.2 Fit With Path Loss Model

The mathematical model from Section 3 does not depend on any assumptions about the spatial continuity of the signal fingerprints. This means we can only model classification accuracy, but not the accuracy of absolute location inference. If the signals *did* display some easily expressible spatial continuity, then a more spatially oriented model would be appropriate. In this section, we show that our experimental WiFi signals do not follow a simple model of

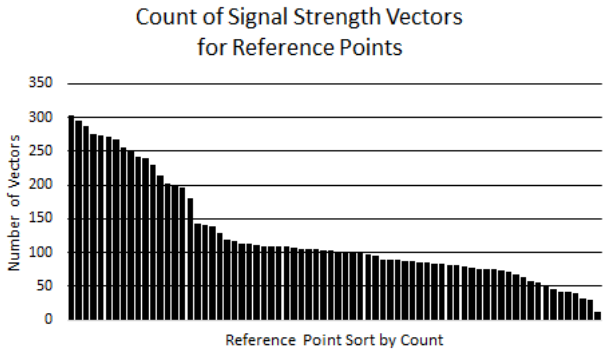

Figure 7: The number of signal strength vectors per reference point varied from 12 to 304, with an average of 125.

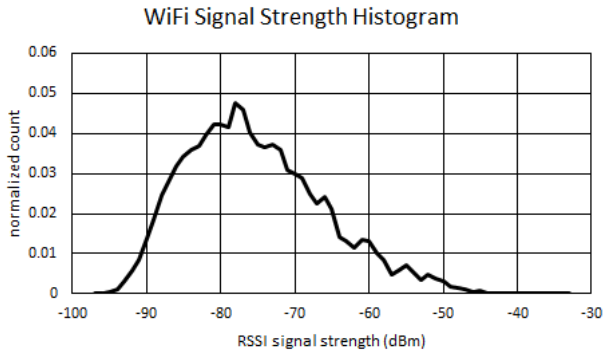

Figure 8: This is a histogram of measured signal strengths, which varied from -97 to -33 dBm.

radio propagation, which helps justify our non-spatial model from Section 3.

One common model for WiFi signal strengths is the path loss model [1, 5]:

$$s = s_0 - 10\eta \log_{10}\left(\frac{d}{d_0}\right) \tag{6}$$

where s is the signal strength at distance d from the base station. The values s_0 and d_0 are a known pair of reference signal strength and distance, respectively. The variable η represents the path loss coefficient. Researchers have investigated more sophisticated path loss models [1, 6] for indoor location, but these require geometric models of the building's walls and/or floors, which are difficult to express in the kind of generalized models we seek.

We fit the path loss model in Equation 6 to data from the three base stations as described in Section 4.1. This meant estimating each base station's (x, y) location and its own values of s_0, d_0, and η. The results for the three base stations are shown in Figure 9. The thick curve in these plots represents the best fit path loss model from Equation 6, and the dots represent the actual signal strengths.

From the plots, it is apparent that the actual signal strengths deviate significantly from the best fit path loss model. For the three base stations, the standard deviations of the error between the actual

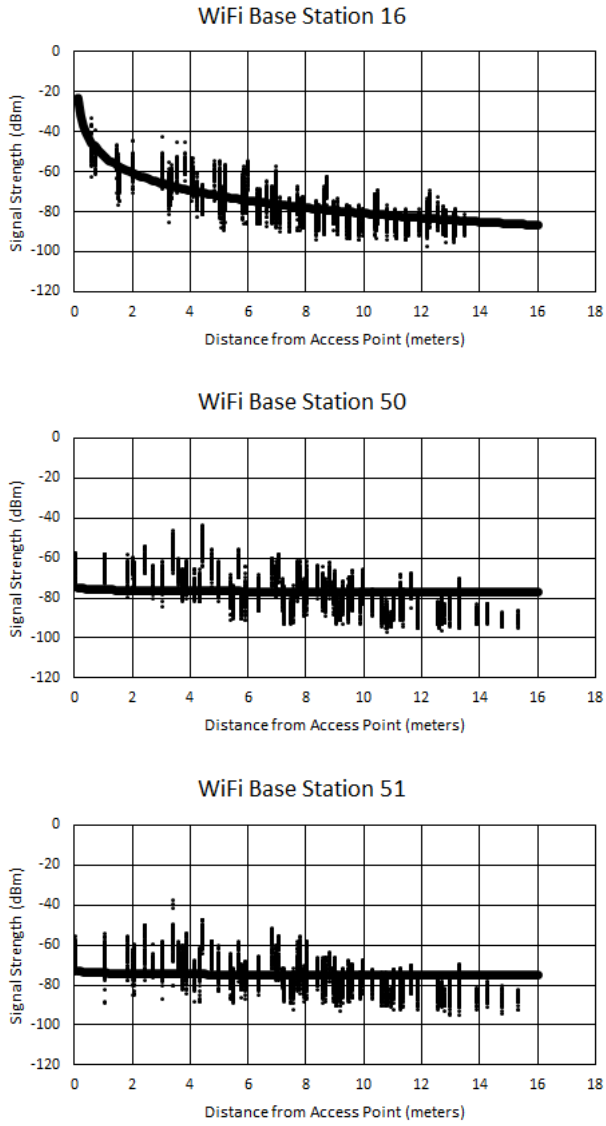

Figure 9: These are best fit curves of the path loss model to WiFi signal strength data.

and best fit models are $(5.72, 9.04, 8.13)$ in dBm. As a comparison, the average standard deviations of the measured signal strengths at the reference points are $(3.64, 3.49, 3.96)$ in dBm, meaning that the modeling error is about $(1.58, 2.59, 2.05)$ times as large as the signal noise, respectively. This indicates that modeling the spatial continuity of signals in this way would introduce significant noise, overshadowing even the inherent measurement noise.

5 VORONOI CELL CLASSIFICATION

In Section 3 we described an idealized system for splitting a region into imaginary, equal-sized cells, each represented by a distinct set of signal strength vectors. This geocoding is aimed at predicting the

performance limits of fingerprint-based location before prototyping or deployment. In other cases, a region may have already been fingerprinted, such as the experimental data described in Section 4. In this postdeployment case, the reference points have already been chosen, and we can compute statistics on the signal strength vectors in the fingerprints. In this section we develop and test a mathematical framework for predicting the performance of a previously fingerprinted space in terms of location cell classification accuracy.

5.1 K Nearest Neighbor Experiment

We begin by assessing the performance of a simple algorithm for determining which region is indicated by a runtime signal strength vector. The kNN algorithm has been used before for WiFi location [1], so we will use it here. The runtime fingerprint is $\mathbf{s}^{(r)}$. kNN compares this to all the calibrated fingerprints from all the reference points. The winning reference point is the one with the most votes out of the top k most similar fingerprints. For the $N = 72$ reference points from the experimental data in Section 4, Figure 10 shows the accuracy results of kNN for $k \in [1, 20]$. The plot shows a wide range of accuracies, with most clustered at relatively low accuracies. The thick black curve is the mean accuracy as a function of k.

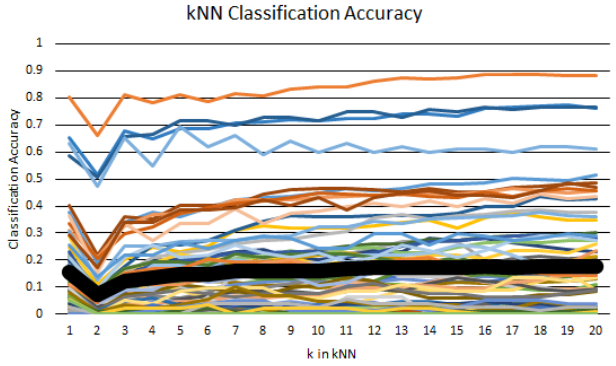

Figure 10: Classification accuracy varies with k in kNN. Each colored curve shows classification accuracy for one of the $N = 72$ reference points. The thick black curve shows the mean accuracy over all reference points.

We computed the results in Figure 10 using a leave-one-out method. For each signal strength vector in the dataset, we found which other vector in the dataset was nearest in signal space. We computed distance in signal space as Euclidean distance: If the runtime fingerprint were $\mathbf{s}^{(r)}$, then the distance to calibrated fingerprint $\mathbf{s}_i^{(c)}$ was $||\mathbf{s}^{(r)} - \mathbf{s}_i^{(c)}||^{1/2}$.

The experimental data from Section 4 was assembled over a period of many months, with signal strength collections happening in bursts over time. During a collection, signal strengths at each reference point and each orientation of the human collector were measured six times in quick succession. It is unrealistic to have calibration fingerprints gathered so near to runtime fingerprints in time. Because of this, we reduced the set of calibration fingerprints to only those that were separated by at least one day from the runtime fingerprint.

5.2 Voronoi Cells

We will model the accuracy result for $k = 1$ in kNN, i.e. the single nearest neighbor case, in order to predict classification accuracy. As before, there are N reference points where the n-dimensional signal strength vectors have been fingerprinted in a calibration phase. In geocoding, the number of reference points was determined by the number of combinations of discrete signal strengths. In this section's analysis, since the calibration phase is already complete at preselected reference points, we instead assume that the signal strength vector $\mathbf{s}_i$ at reference point i is normally distributed, i.e. $\mathbf{s}_i \sim \mathcal{N}(\mu_i, \Sigma_i)$.

Upon receiving a runtime signal strength measurement $\mathbf{s}^{(r)}$, the system compares $\mathbf{s}^{(r)}$ to each mean vector μ_i of the N reference points. This is slightly different from kNN where *each* calibrated fingerprint is compared to the runtime vector. Instead, we are modeling a comparison to each reference point's *mean* vector as an approximation to kNN. The winning reference point i^* is the one that is nearest $\mathbf{s}^{(r)}$, i.e. $||\mu_{i^*} - \mathbf{s}^{(r)}|| < ||\mu_i - \mathbf{s}^{(r)}|| \; \forall \, i \in \{1 \ldots N\}, i \neq i^*$. We can think of a Voronoi diagram in signal space, where the i^{th} Voronoi cell $\mathcal{R}_i$ contains the signal vectors that are nearest μ_i. The distance metric $||\cdot||$ could be any metric, such as Euclidian, Mahalnobis, or Manhattan. We used Euclidean distance for our experiments.

The signal noise Σ_i will cause errors in determining the winning reference point. If the runtime signal $\mathbf{s}^{(r)}$ is drawn from Voronoi cell $\mathcal{R}_i$, then $\mathbf{s}^{(r)} \sim \mathcal{N}(\mu_i, \Sigma_i)$. The probability of $\mathbf{s}^{(r)}$ being nearest μ_i is

$$P(\mathbf{s}^{(r)} \in \mathcal{R}_i) = \int_{\mathcal{R}_i} g(\mathbf{s}; \mu_i, \Sigma_i) d\mathbf{s} \tag{7}$$

where $g(\mathbf{s}; \mu_i, \Sigma_i)$ represents the multivariate normal distribution:

$$g(\mathbf{x}; \mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu)\right)$$

Note that the regions $\mathcal{R}_i$ are regions in signal space $\mathbf{s}$, not regions in the indoor space where we are trying to measure location. The integral is taken over the continuous space of signal vectors in signal space region $\mathcal{R}_i$. Also, Σ is a diagonal matrix, because we consider signal values as statistically independent. In our experiment, we computed an individual variance for each base station at each reference point, i.e. $\sigma_{i,j}^{(r)}$ is different for each i, j, where i indexes the $N = 72$ reference points and j indexes the $n = 3$ base stations.

5.3 Experimental Results

Section 5.1 describes how we computed the experimental accuracies, of which we only used $k = 1$ for comparing to the model.

We used Equation 7 to predict the winning reference point for each possible runtime signal $\mathbf{s}^{(r)}$ in the full space of signals $\mathbf{s}$. These integrals were computed numerically.

For each of the $N = 72$ reference points, the vertical axis of Figure 11 shows the predicted probability of the correct reference point being chosen based on Equation 7. The horizontal axis shows the

$k = 1$ rate of choosing the correct reference point in our experiment with WiFi signal strengths. Ideally these should be equal for all the reference points. We do see the results clustered around a diagonal, indicating good agreement. Averaged over all reference points, the model predicts an average correct classification rate of 0.181, while the actual correct classification rate was 0.154, a difference of 2.7 percentage points on a $[0, 1]$ scale. Thus our model does a good job predicting overall location accuracy.

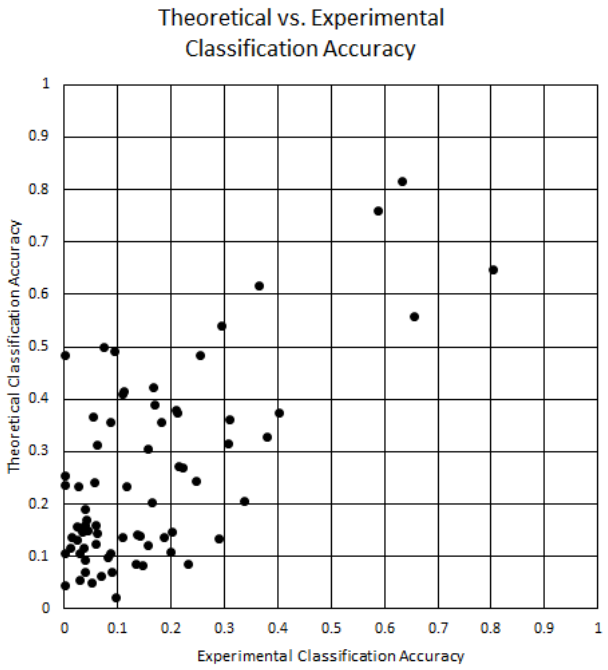

Figure 11: This shows the classification accuracy for each of the $N = 72$ reference points as predicted by our model (vertical axis) and realized with a nearest neighbor algorithm (horizontal axis).

6 CONCLUSIONS AND FUTURE WORK

We presented a predeployment and postdeployment model for estimating the accuracy of location-based fingerprinting. The predeployment model includes effects of signal noise, signal quantization, spatial quantization, and calibration effort. We showed examples of how the model responds to changes in these parameters. It introduces the concept of noisy reverse geocoding. The postdeployment model gives accuracy estimates for an existing fingerprint installation, and we verified its accuracy with WiFi data. These types of models are important for anticipating the accuracy of location-based fingerprinting to make decisions about deployments.

Future work should consider more sophisticated models of the ambient signals. For instance, we know that extreme signal values, such as extreme temperatures and radio signal strengths, are relatively rare compared to signal values in the mid-range. Modeling a probability distribution of signal values would likely affect the predicted location accuracies. We also know that signals from nearby locations are likely correlated with each other, because the ambient signals we can measure often vary smoothly over space. Modeling this spatial correlation may be a way to predict absolute location errors in addition to the classification errors that we currently model.

REFERENCES

[1] Paramvir Bahl and Venkata N Padmanabhan. 2000. RADAR: An in-building RF-based user location and tracking system. In *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No. 00CH37064)*, Vol. 2. Ieee, 775–784.

[2] C. BASRI and A. El Khadimi. 2016. Survey on indoor localization system and recent advances of WIFI fingerprinting technique. In *2016 5th International Conference on Multimedia Computing and Systems (ICMCS)*. 253–259.

[3] Roberto Battiti, Mauro Brunato, and Andrea Delai. 2003. *Optimal wireless access point placement for location-dependent services.* Technical Report. University of Trento.

[4] Manuel Blanco-Muriel, Diego C Alarcón-Padilla, Teodoro López-Moratalla, and Martín Lara-Coira. 2001. Computing the solar vector. *Solar energy* 70, 5 (2001), 431–441.

[5] Atreyi Bose and Chuan Heng Foh. 2007. A practical path loss model for indoor WiFi positioning enhancement. In *2007 6th International Conference on Information, Communications & Signal Processing*. IEEE, 1–5.

[6] Giuseppe Caso and Luca De Nardis. 2015. On the applicability of multi-wall multi-floor propagation models to wifi fingerprinting indoor positioning. In *Future Access Enablers of Ubiquitous and Intelligent Infrastructures*. Springer, 166–172.

[7] Christian Hirt, Sten Claessens, Thomas Fecher, Michael Kuhn, Roland Pail, and Moritz Rexer. 2013. New ultrahigh-resolution picture of Earth's gravity field. *Geophysical research letters* 40, 16 (2013), 4279–4283.

[8] Kamol Kaemarungsi and Prashant Krishnamurthy. 2004. Modeling of indoor positioning systems based on location fingerprinting. In *Ieee Infocom 2004*, Vol. 2. IEEE, 1012–1022.

[9] John Krumm, Gerry Cermak, and Eric Horvitz. 2003. Rightspot: A novel sense of location for a smart personal object. In *International Conference on Ubiquitous Computing*. Springer, 36–43.

[10] Germán Martín Mendoza-Silva, Philipp Richter, Joaquín Torres-Sospedra, Elena Simona Lohan, and Joaquín Huerta. 2018. Long-term WiFi fingerprinting dataset for research on robust indoor positioning. *Data* 3, 1 (2018), 3.

[11] Kasandra O'Malia. 2013. Thermistors As Accurate Temperature Sensors Part 1: Introduction and Methods. https://www.fierceelectronics.com/components/thermistors-as-accurate-temperature-sensors-part-1-introduction-and-methods. [Online; accessed 16-June-2020].

[12] World Meteorological Organization. 2020. *World Meteorological Organization Global Weather Climate Extremes Archive.*

[13] Asim Smailagic, Jason Small, and Daniel P Siewiorek. 2000. Determining user location for context aware computing through the use of a wireless LAN infrastructure. *Institute for Complex Engineered Systems Carnegie Mellon University, Pittsburgh, PA* 15213 (2000).

[14] I Smith, K Tang, T Sohn, F Potter, A LaMarca, Jeffrey Hightower, D Haehnel, J Froehlich, E de Lara, MY Chen, et al. 2005. Are GSM phones THE solution for localization?. In *Seventh IEEE Workshop on Mobile Computing Systems & Applications (WMCSA'06 Supplement)*. IEEE, 34–42.

[15] Wikipedia contributors. 2020. What3words — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=What3words&oldid=960223865. [Online; accessed 14-June-2020].