EvaLoc: Evaluating Performance Degradation in Wireless Fingerprint-based Indoor Localization

Hande Hong
School of Computing
National University of Singapore
honghand@comp.nus.edu.sg

Paramasiven Appavoo
School of Computing
National University of Singapore
pappavoo@comp.nus.edu.sg

Chengwen Luo
College of Computer Science and Software Engineering
Shenzhen University
chengwen@szu.edu.cn

Mun Choon Chan
School of Computing
National University of Singapore
chanmc@comp.nus.edu.sg

ABSTRACT

Many WiFi fingerprint-based indoor localization approaches have been proposed to ease deployment and minimize infrastructure requirement. While researchers have devoted extensive efforts to improving the accuracy of these approaches, the user experience of such deployments in practice is typically far below expectation. One reason that contributes to such discrepancy is that while researchers often evaluate their systems in stable and “benign” environments, the actual environments can be much more dynamic and noisy. In this paper, we address this issue in the following manner. First, we identify factors that can result in significant degradation of localization performance and explore how these factors can be modeled in the localization process. Next, we design a system, EvaLoc, that takes fingerprinting data collected as input and provides accuracy prediction on the localization performance under different conditions. Our evaluation in 15 different locations covering around 25000 m² shows that EvaLoc is able to produce localization result that better matches the user experience.

CCS CONCEPTS

• Human-centered computing → Smartphones; Ubiquitous and mobile computing design and evaluation methods; • Computing methodologies → Simulation evaluation; • Networks → Wireless access points, base stations and infrastructure;

KEYWORDS

Indoor localization, fingerprinting, simulation, error estimation

1 INTRODUCTION

WiFi fingerprint-based indoor localization has attracted extensive efforts from researchers due to its ease of deployment and low infrastructure requirements. While researchers have proposed various kinds of fingerprint-based indoor localization methods with accuracies ranging from 1m to 5m [1, 7, 8, 11, 12, 14, 16, 18–20], indoor localization techniques are still underutilized in daily life and location-based application. Since indoor localization is a key enabling technology for many applications such as indoor navigation[5, 22], smart homes[4], and targeted advertisement[10], this lack of deployment is somewhat of a surprise. We believe that one of the issues limiting the use of WiFi fingerprint-based techniques is the big performance gap between laboratory results and practical usage.

The differences between research prototypes and practical deployments are mainly caused by the differences in behavior between researchers and users. Much research measures the fingerprints under favorable and stable settings, for example, by assuming that there is limited human movement in the environment. While such assumption help to ensure that the results are consistent and repeatable, the usage and environment seen by users can be very different. In practice, a user can face different directions and the amount of human activity in the surroundings can change. Understanding the impact of such factors on localization accuracy will help bridge the research and practice gap.

Despite extensive research on improving WiFi fingerprinting localization accuracy[1, 7, 14, 20] and reducing the calibration effort[8, 11, 12, 16, 18, 19], there is surprisingly little systematic study on the factors that will impact the accuracy. In this work, we try to answer the following questions. (1) What is the impact of user behavior and environmental noise on WiFi RSS; (2) How can the impact of these factors be accessed in the localization system?

The contributions of this paper are as follow:

• We analyze the impact of human blocking, phone placement, and different level of human mobility (crowdedness) on the WiFi fingerprint detected. We also present models to explain and predict how these factors affect the Receiver Signal Strength (RSS).
We design EvaLoc, a tool that takes the fingerprinting data collected as input and provides accuracy prediction on the localization performance achievable under different conditions. EvaLoc treats the fingerprint data and localization algorithm independently and can also be used to access either the quality of the fingerprint data or the performance of the localization algorithm.

Our evaluation in 15 different locations covering 25000 m² shows that EvaLoc is able to produce localization result that better matches the user experience.

The rest of the paper is organized as follows. In section 2, we study how RSS will be affected under different conditions. We model signal blocking effect under different condition in section 3. We present the design of EvaLoc in section 4 and the evaluation result in section 5. Finally, we discuss related work and summarize our work in section 6 and 7 respectively.

Figure 1: Human body blocking effect experiment when the user walks pass an AP in a corridor

2 FACTORS AFFECTING RSS SIGNAL COLLECTED

In this section, we provide measurements to illustrate three factors that can have significant impacts on the RSS of WiFi signals collected.

2.1 Body Blocking Effect

The most common source of signal blocking comes from the user himself. In order to study the blocking effect of the human body, we measure how the signal strength varies when the position and orientation of the user changes. The setup of the experiments is shown in Figure 1. In the experiment, the user holds the phone in front of his body and rotated his body slowly through 360 degrees in each position. A full rotation lasts for 2 minutes. We group the 360-degree movement into discrete steps and we obtained about 80 samples in each position. As shown in Figure 1(a), the user moved his position closer to the access point and then moved away. We plot the RSS trend in Figure 2(a)-(d) with user initially facing north at 4 different locations (A, B, C, and D). We repeated the experiment without human blocking and also plot the RSS trend in Figure 2(a)-(d) as dotted lines.

From Figure 2(a)-(d), we can observe low RSS valley when a user’s back is facing the access point from angles 260 to 350 degrees in positions A and B. When the user walks past the access point to positions C and D, the signal valley occurs at angles from 50 to 140 degrees. Because of signal fluctuation and uneven rotation speed, there is a slight shift in the position of the valley. The signal valley shows a trapezoidal groove with a maximum RSS drop ranging from 10-15 dBm.

We did similar experiments in the non-line-of-sight scenarios and still observed significant body shielding effect. The human body blocking phenomenon generally exists when a user holds the phone in his hand and has his back facing the corresponding access point.

2.2 Device in Pockets and Bags

When users do not hold their phones in their hands, they usually put them in pockets or bags. Many applications can still function well even when phones are placed inside pockets or bags. For example, a user can enable voice navigation even when the phone is in a pocket. To understand how different kinds of pockets and bags have an impact on RSS, we collected signal profiles when the phone is placed in a cotton pocket, a nylon bag, jeans pocket, and no pocket. To our surprise, we see no obvious change in signal strength when phones are put into pockets or bags in comparison with putting in the users’ hand. The major impact on signal blocking when phones are put into the pocket still comes from the human body itself.

2.3 Crowded Environment

The two factors mentioned above are affected by the user’s behavior. Another important factor that needs to be taken into account is when the environment is highly dynamic with people walking around. Due to movement in the vicinity, signal propagation can be blocked from time to time by a single person or multiple people. In order to understand the impact of crowded environments, we collect WiFi beacon data and crowd information in a campus canteen, which has a highly dynamic crowd. As implementing a manual count of people can be quite labor-intensive, we use the method mentioned in [3] that scans the probe request frames sent by smartphones to estimate the number of people. At the same time, we use smartphones to collect beacon data continuously for more than seven hours. The result is plotted in Figure 3 and Figure 4.

Figure 3 shows the signal strength collected by the mobile phones from beacons emitted by the same AP. As we can observe, RSS decreases with larger crowd size. Figure 4 shows how the standard deviation of RSS varies with crowd size in the same settings. The impact on signal variation is also very clear. There is a clear relationship between these two parameters. Typically, in most published work, evaluation of a localization system is done when there is no crowd or very limited human movement.

3 MODELING SIGNAL BLOCKING EFFECT

3.1 Modeling Body Blocking Effect

In this paper, we discuss the most common position to hold the phone as shown in Figure 5(a). Signals from the back of the body are blocked. To correctly understand the impact of body shielding on RSS, we need to derive each component of the body shielding effect, namely RSS blocking angle, RSS blocking orientation and RSS drop magnitude. We present how we derive all the related information as follow:

**RSS blocking angle:** We apply the body shielding model from [21] as shown in Figure 5. While there are multiple gestures that
Figure 2: A to D represents the position where the signal profile is collected in Figure 1. A clear low RSS valley can be observed when human body blocks the direct sight of mobile phone and access point. We can observe a shift of low RSS valley when user walk passes the access point from A to D.

RSS blocking angle: Since the users’ devices will suffer from body shielding signal drop only when their users turn their back to the AP. Thus by inferring the direction of the AP relative to the user, we are able to infer the RSS blocking area when incorporating with the above RSS blocking angle. As shown in Figure 6(a), we estimate the direction of the AP based on the spatial distribution of RSS. At each of the locations drawn as a circle, we use the direction from the current position pointing to the strongest RSS position indicated by the black arrow in Figure 6(a).

Once we know the AP direction and the RSS blocking angle, we can identify the RSS drop area. We mark the RSS drop area in Figure 6(a) with a red color quarter circle. When users walk towards the direction within this red quarter with their phone holding in front of their chests, the signal received by the mobile device will have a significant drop because of blocking effect.

RSS drop magnitude: The maximum RSS drop as shown in Figure 6(b) happens around the opposite direction of the AP direction. To infer this value, we make use of the noise contained in the raw database. In most of the cases, data collected by users may already contain the body blocking effect. Thus during data pre-processing, we not only filter out noise data but also record the maximum RSS drop ratio for each AP in each location. In cases where we cannot obtain good estimates for the RSS drop magnitude, we use the average from neighboring locations which should have similar body blocking impact.

We identify the RSS drop value based on the relative position of the walking direction in the red quarter area as shown in Figure 6(b). We model the signal drop as inverted isosceles trapezoid and use Equation 1 and 2 to derive RSS drop value:

$$\Delta = \frac{\Theta_{ap} - \Theta_{user}}{\Delta_{drop}/2}$$

where \( \Delta_{RSSbody} \) is the RSS drop value, which is decided by the RSS drop ratio \( \tau \) and maximum RSS drop \( \Delta_{Max} \). \( \Theta_{ap} \) is the opposite direction of the AP direction \( \Theta_{ap} \). \( \Theta_{user} \) is the user’s walking direction within this red quarter area.

$$\Delta_{RSS} = \begin{cases} \Delta_{Max} & \text{if } \tau \leq 1/2 \\ \Delta_{Max} \times (1-\tau) & \text{if } 1/2 \leq \tau \leq 1 \\ 0 & \text{otherwise} \end{cases}$$
direction. $\theta_{drop}$ is the RSS drop angle, in this case, we estimate it as 90 degrees. For example, in Figure 6(b), we assume the AP direction is at 180 degrees and the user walks through different directions labeled as 1 to 6. Based on Equation 2, walking directions 1, 2, 5 and 6 do not incur an RSS drop. Walking directions 3 and 4 incur an RSS drop equal to $\Delta_{Max}/2$ and $\Delta_{Max}$ respectively.

### 3.2 Modeling Pocket Effect

As mentioned in section 2.2, the pocket material does not have a significant impact on signal strength level. Thus, most of the effects come from the human body. We can still use Equation 1 and 2 to model the impact, but adjust the walking direction as the facing orientation of the pocket. For example, when we put the phone in the front shirt pocket, $p$, the distance between body and phone is close to 0 which indicates that $\theta_{drop}$ is estimated as $180^\circ$. When we put the phone in the side pocket, the body width $b$ is smaller as the side face of the body. Compared to the front of the body, the side of our body does not block as much signal. Thus it only creates a narrow area band where signal will be blocked. Such estimation is proved by experiment in section 2.2. We still use Equation 2 to calculate RSS drop $\Delta_{RSS}$, but adjust the RSS drop angle $\theta_{drop}$ empirically, to be smaller as 45 degrees.

### 3.3 Modeling Crowd Density

With people walking around and blocking signals from different sources, the signal detected will be much more dynamic. To gauge the impact, we divide crowd level into 3 coarse categories, namely dense ($> 0.5\text{person}/m^2$), light ($> 0.05\text{person}/m^2$ but $< 0.5\text{person}/m^2$) and sparse ($< 0.05\text{person}/m^2$). We adopt the method in [3] for its ease of implementation. WiFi monitors include Raspberry Pi 3 with one D-Link wireless USB adapters (DWA-132) are used to capture the Wifi beacons needed for the estimated count.

In order to quantify the impact of crowd, we look at the difference between RSS detected from different crowd levels. Figure 7 shows the drop in signal strength for both dense and light crowd as compared to sparse crowd level. We can see that the light crowd environment incurs a slight decrease in RSS while in the area with dense crowd, RSS drop can be as large as 10.5 dBm. We also plot the RSS standard deviation with respect to RSS level. In Figure 8, we see a clear trend of increase in signal variation with increasing RSS. Again, the variation for dense crowd is much more than environment with light crowd.

In order to model crowd effect, we first use linear regression to derive the (linear) relationships between RSS drop/variation and...
RSS level for both dense and light crowd environments based on the data collected. We then estimate the crowd level (dense, light or sparse) based on crowd density. With the crowd level estimate, the appropriate linear relationships are chosen to derive the statistical parameters (mean and standard deviation) to be used in the EvaLoc simulation introduced in the next section.

4 OVERVIEW OF EVALOC

The architecture of EvaLoc is shown in Figure 9. EvaLoc has three main components:

(1) **Data Preprocessing**. A data cleaning process is performed to remove the effect of blocking by human body (Section 4.1). We also record down the maximum RSS drop magnitude for each AP in each location.

(2) **Fingerprint Generation**. We build a Gaussian distribution for each of the selected BSSIDs in different locations. Then based on the distribution, we generate fingerprint samples that take into account different signal blocking effects (Section 3).

(3) **Simulation**. Finally, given the localization algorithm, EvaLoc predicts the accuracy of the target indoor location (Section 4.3).

In the rest of this section, we will introduce each component of the system in detail.

4.1 Data Preprocessing

Data Preprocessing consists two steps. In the first step, outliers are removed with density based clustering [2]. As for body blocking effect, we notice in Figure 2 that signal profile will drop a few dBm which separate them from the normal signal without body blocking effect. Thus after removing outliers, EvaLoc uses k-mean clustering to divide the signal data. We will remove the cluster with obvious lower signal value and keep the higher RSS values as the training data for fingerprint sampling.

4.2 Fingerprint Generation

With the radio data, we can derive the signal distribution of RSS in each location either using a Gaussian distribution or histogram-based estimation [20] for all the BSSIDs. In this work, we use Gaussian-based sampling for ease of implementation and less storage cost. Combined with response rate for each of the BSSIDs, we can generate virtual fingerprints. In Figure 10(b), we show the fingerprints sampled for 62 BSSIDs. Compared to the ground truth samples in Figure 10(a), we can see that Gaussian-based sampling match the ground truth sampling well.

In order to include the impact from body blocking and crowd density, we revise the Gaussian-based sampling to get more precise samples that are closer to ground truth condition. We present our refined Gaussian-based fingerprint sampling in Algorithm 1.

![Figure 9: Overview of EvaLoc](image)

**Algorithm 1: Fingerprint Sample Algorithm**

\[
\text{Input: Location } L, \text{ mean } \nu, \text{ variation } \sigma_L^2, \text{ crowdedness level } I_{\text{crowd}}, \text{ user facing direction } \Theta_{\text{user}} \\
\text{Output: Fingerprint Sample } f \cdot p \\
// Get RSS and its deviation adjustment based user facing direction and environment crowdedness \\
\text{ARSS}_{\text{bod}} = \text{getBodyRSSDrop}(\Theta_{\text{user}}); \\
\text{ARSS}_{\text{crowd}} = \text{getCrowdRSSDrop}(I_{\text{crowd}}); \\
\text{ARSS}_{\text{etCrowdV ar Increment}} = \text{getCrowdV arIncrement}(I_{\text{crowd}}); \\
\sigma_L^2 = \sigma_L^2 + \text{ARSS}_{\text{etCrowdV ar Increment}}; \\
\nu = \nu + \text{ARSS}_{\text{bod}} + \text{ARSS}_{\text{crowd}}; \\
N = \text{The No. of BSSID detected in location } L; \\
\text{for } i = 1 : N \text{ do} \\
\text{with response rate } p_i \\
\text{set } r_i = \text{GaussianSample}(\nu, \sigma_L^2); \\
\text{end} \\
f \cdot p = \{r_1, r_2, \ldots, r_m\}, m <= N \\
\text{return } f \cdot p;
\]

In each location \(L\), we derive the Gaussian parameters, mean \(\nu\) and variation \(\sigma_L^2\), based on the radio map data. After that, we adjust these two parameters to include different kinds of impact from body blocking or crowd density. In the end, we can use Gaussian sampling to generate RSS for each AP with a specific response rate. Combining RSS from different APs, we can generate a fingerprint sample for each location. Based on this sampling method, we perform indoor localization simulation in next section.

4.3 Indoor Localization Error Estimation

With the above refined fingerprint sampling tool, we can simulate the localization error in Algorithm 2. Given location \(L\), we can easily generate a virtual fingerprint. On inputting this fingerprint to any localization algorithm that we want to test, we can get a suggested location. By calculating the distance between this suggested location and \(L\), we get a location error estimation, \(e_iL_i\). We run the simulation thousands of runs until the localization error converges to a stable value.

5 EVALUATION

The evaluation is separated into two parts. In the first part, we evaluate EvaLoc in two different environments, namely an office floor and within a laboratory, as shown in Figure 11. The data collection and accuracy estimation were performed by users who are quite familiar with indoor localization technique. With this setting, we can do a detailed evaluation for each component of EvaLoc.
Algorithm 2: Error Estimation Algorithm

1. **Input**: Location $L$, Localization algorithm $\Gamma$
2. **Output**: Localization error $E(L)$
3. $K$ is the number of samples we generated for location $L$;
4. for $i = 1 : K$ do
5.  Get fingerprint sample
6.  from Refined Gaussian-based sampling;
7.  calculate $e(L_i)$ for location $L$ with algorithm $\Gamma$;
8.  add to $E(x)$;
9. end
10. $E(x)/K = e$
11. return $E(x)$;

In the second part, we release our indoor localization mobile APP including data collection, indoor localization tool and evaluation service for use by users who have no prior experience with EvaLoc. We obtained data for 10 shopping malls, an exhibition site, a food court and a museum that were used in the evaluation.

5.1 Implementation

In the first part of the evaluation, the accessible area in Figure 11 is divided into square cells automatically for the ease of data processing. The black dots indicate the location of the cell center. In the small room, the average distance between neighboring cell centers is 0.87 meter, while in the office floor the average cell length is 3.5 meter. The data was collected with three different phone models: Nexus 5, Motorola Nexus 6 and Galaxy S4. Users collected data by walking around the site and uploading their walking trajectories similar to the approaches used in [8, 12].

As the fingerprints are collected while walking and not through tedious cell-by-cell measurements, it took less than an hour to perform data collection. We collected an average of 25 samples from each cell for the 91 cells on the office floor. Note that, as our main objective is to understand how different factors affect the accuracy of WiFi fingerprint localization systems, collecting data quickly, even if the data collected is noisy, is sufficient for our purpose. However, in order to do ground truth verification in section 5.4, we need to do real localization under different settings using traditional site survey methods to estimate the errors which took more than 4 hours.

Figure 10: WiFi Signal sampling based on Gaussian distribution correctly match ground truth

![Figure 10: WiFi Signal sampling based on Gaussian distribution correctly match ground truth](image)

![Figure 11: Floor plan for two different environments we evaluate. Arrows give the AP direction estimation using EvaLoc. Noted not all the cells are plotted for clear illustration. Blue AP icon marks AP ground truth location.](image)
When the body rotates to angle around 180 degrees, major BSSIDs are blocked, it has less impact on the localization performance. Compared to those, noisy environments with human walking around frequently by a body and the phone in the side pocket increase the localization cell length for all the location. We can also see that signal blocked away to 5 cells away. This confirms our daily experience that localization errors in each cell and show the result in Figure 13.

Converges to a stable value. We calculate the distribution of average thousand times until the localization error calculated by algorithm 2 side pocket, light crowd, and dense crowd. The effect of these factors conditions: No impact, includes body blocking effect, phone in the side pocket, and environment noise to localization performance.

In Figure 13, we compare the localization errors under five different benchmark. In the following sections, we evaluate the impact from body blocking, pocket and environment noise to localization performance based on Algorithm 2.

Figure 12 plots the localization errors using different algorithms: Nearest Neighbour (Radar[1]), Weighted Nearest Neighbour (PiLoc[8]) and Bayesian Inference (Horus[20]) for the same location. We simulate the impact of the body shielding effect with the user in different orientations. Since WiFi signals comes from all directions, any orientation of the body can block some sets of APs. The blocking of received signals can have a diverse impact on the final localization errors as shown in Figure 6. From Figure 12, we can observe the fluctuation of localization accuracy under the body blocking effect. When the body rotates to angle around 180 degrees, major BSSIDs that have stronger RSS are blocked. In this case, we see a significant increase in the localization error. The increase in errors can be as large as 1 meter. This figure also shows that Horus and PiLoc has a better localization performance than Radar. Thus in the later evaluation, we use the average error of these two methods as the benchmark.

5.2 Body Facing Orientation

In the following sections, we evaluate the impact from body blocking, pocket and environment noise to localization performance based on Algorithm 2.

Figure 12 plots the localization errors using different algorithms: Nearest Neighbour (Radar[1]), Weighted Nearest Neighbour (PiLoc[8]) and Bayesian Inference (Horus[20]) for the same location. We simulate the impact of the body shielding effect with the user in different orientations. Since WiFi signals comes from all directions, any orientation of the body can block some sets of APs. The blocking of received signals can have a diverse impact on the final localization errors as shown in Figure 6. From Figure 12, we can observe the fluctuation of localization accuracy under the body blocking effect. When the body rotates to angle around 180 degrees, major BSSIDs that have stronger RSS are blocked. In this case, we see a significant increase in the localization error. The increase in errors can be as large as 1 meter. This figure also shows that Horus and PiLoc has a better localization performance than Radar. Thus in the later evaluation, we use the average error of these two methods as the benchmark.

5.3 Cell Localization Error Simulation

In Figure 13, we compare the localization errors under five different conditions: No impact, includes body blocking effect, phone in the side pocket, light crowd, and dense crowd. The effect of these factors is modeled based on the measurement data presented in Section 3.

In each cell on each floor, we simulate localization more than a thousand times until the localization error calculated by algorithm 2 converges to a stable value. We calculate the distribution of average errors in each cell and show the result in Figure 13.

The distribution curves tell us that different parts of a single floor can have quite different localization accuracy ranging from 1 cell away to 5 cells away. This confirms our daily experience that localization performance is not stable when we move around. In both of the figures, localization achieved the best performance within three cell length for all the location. We can also see that signal blocked by a body and the phone in the side pocket increase the localization error slightly. Since a side pocket only triggers a small blocking angle, it has less impact on the localization performance. Compared to those, noisy environments with human walking around frequently can deteriorate localization performance to a greater extent which may make the indoor localization system unusable.

Besides the overall localization error, we also found that body shielding has a different impact on different locations on the same floor. To better understand this phenomenon, we plot the localization error heatmap in Figure 14 and Figure 15. Darker color means higher localization error. In the “No Impact” case, most locations show light color. With body blocking impact, errors in some of the locations become extremely high. Such locations are marked by small arrows. Lightly crowded environments increase the error slightly across all the locations, while dense crowd condition can lead to poor localization performance.

5.4 Effectiveness of EvaLoc

To verify the effectiveness of EvaLoc, we compare the output of EvaLoc with the “ground truth” by performing actual localization in different locations. In each of the locations, we perform evaluation under two different scenarios. In the first scenario, evaluation is performed assuming that the room is empty and there is no impact from human blocking effect. This matches the “No Impact” case in Figure 13. In the second scenario, the user can change his orientation. only light crowd in the surrounding is considered. Such a scenario is closer to the typical usage environment.

We plotted the ground truth(GT) localization performance under these two conditions as solid shapes connected by lines in Figure 16. For ease of comparison, we also plot the simulation result from EvaLoc as hollow shapes connected lines. We can observe that the localization accuracy between the “No Impact” scenario and the “Light Crowd + Body” scenario is significantly different. On the other hand, the output of EvaLoc in the “Light Crowd + Body” scenario is much closer to the ground truth.

The ground truth localization error result also shows higher error than the simulation result under impact from body effect and crowded environment. The higher error may have come from the fact that we do not have enough well calibrated WiFi fingerprints. In the simulation procedure, we can generate enough samples to make sure that the error converges. In the actual localization, due to the time limitation, the amount of data collected may be insufficient for a large floor area. Nevertheless, it is clear that incorporating factors such as signal blocking and crowd significantly improve the accuracy of the estimation compare to the baseline of assuming there is no such impact.

Overall, the results show that EvaLoc is able to provide a much better estimation of a system’s accuracy in realistic environments and usage scenarios. Considering the long duration (260 min) it took to verify the localization accuracy, EvaLoc also improves the localization evaluation efficiency.

5.5 Evaluation with More Users

The above evaluation was performed by researchers who are familiar with EvaLoc. To gauge EvaLoc’s performance with general users, we make EvaLoc available for general use. Over two months, data were collected by different users from more than 20 locations including shopping malls, museum and food courts. The users also have the option to indicate their actual location so that the localization error can be measured as feedback to improve the system.

5.5 Evaluation with More Users

The above evaluation was performed by researchers who are familiar with EvaLoc. To gauge EvaLoc’s performance with general users, we make EvaLoc available for general use. Over two months, data were collected by different users from more than 20 locations including shopping malls, museum and food courts. The users also have the option to indicate their actual location so that the localization error can be measured as feedback to improve the system.
Figure 13: Localization performance under different conditions in different environments

(a) Localization in small lab
(b) Localization in large office floor

Figure 14: Error heatmaps for a small lab

(a) No impact
(b) Body Blocking
(c) light crowd
(d) dense crowd

Figure 15: Error heatmaps for a large office floor

Figure 16: Localization performance evaluated by EvaLoc and ground truth under three conditions. The big performance gap between “No Impact” and “Light Crowd + Body” reflects the differences between research prototypes and practical deployment.

Table 1 shows the results for the measurements in 13 of the locations. The other locations were not included because the fingerprint data was too sparse. The “Steps” column indicates the effort spent on collecting data with each step annotated with one fingerprint sample. The Error(user) column shows the location errors computed from the user indicated ground truth and the Error(No
Table 1: Using EvaLoc for indoor localization evaluation

<table>
<thead>
<tr>
<th>Location</th>
<th>Steps</th>
<th>Cell Num</th>
<th>Aver. Cell Length</th>
<th>Steps/Cell</th>
<th>Error(No impact)</th>
<th>Error(EvaLoc)</th>
<th>Error(User)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shopping Mall A</td>
<td>603</td>
<td>107</td>
<td>2.4</td>
<td>5.6</td>
<td>1.7</td>
<td>2.9</td>
<td>3.0</td>
</tr>
<tr>
<td>Shopping Mall B</td>
<td>1434</td>
<td>108</td>
<td>2.9</td>
<td>13.2</td>
<td>3.2</td>
<td>3.7</td>
<td>4.0</td>
</tr>
<tr>
<td>Exhibition Site</td>
<td>5926</td>
<td>304</td>
<td>4.5</td>
<td>19.4</td>
<td>3.2</td>
<td>4.1</td>
<td>4.0</td>
</tr>
<tr>
<td>Shopping Mall D</td>
<td>415</td>
<td>99</td>
<td>3.3</td>
<td>4.2</td>
<td>3.5</td>
<td>4.0</td>
<td>4.5</td>
</tr>
<tr>
<td>Shopping Mall E</td>
<td>992</td>
<td>55</td>
<td>6</td>
<td>18.0</td>
<td>3.5</td>
<td>6.4</td>
<td>4.5</td>
</tr>
<tr>
<td>Shopping Mall F</td>
<td>654</td>
<td>27</td>
<td>5</td>
<td>24.2</td>
<td>3.7</td>
<td>6.1</td>
<td>5.0</td>
</tr>
<tr>
<td>Shopping Mall G</td>
<td>2809</td>
<td>48</td>
<td>3.5</td>
<td>58.5</td>
<td>3.5</td>
<td>4.8</td>
<td>5.5</td>
</tr>
<tr>
<td>Museum H</td>
<td>1164</td>
<td>71</td>
<td>5.2</td>
<td>16.4</td>
<td>4.8</td>
<td>6.7</td>
<td>6.0</td>
</tr>
<tr>
<td>Shopping Mall I</td>
<td>561</td>
<td>42</td>
<td>4.3</td>
<td>13.3</td>
<td>4.5</td>
<td>7.4</td>
<td>6.0</td>
</tr>
<tr>
<td>Shopping Mall J</td>
<td>2853</td>
<td>85</td>
<td>6</td>
<td>33.5</td>
<td>4.5</td>
<td>7.6</td>
<td>6.0</td>
</tr>
<tr>
<td>Shopping Mall K</td>
<td>704</td>
<td>47</td>
<td>5.1</td>
<td>14.9</td>
<td>3.4</td>
<td>5.4</td>
<td>6.0</td>
</tr>
<tr>
<td>Food Court L</td>
<td>1072</td>
<td>71</td>
<td>4.9</td>
<td>15.1</td>
<td>4.8</td>
<td>6.8</td>
<td>7.5</td>
</tr>
<tr>
<td>Shopping Mall M</td>
<td>4585</td>
<td>117</td>
<td>5.8</td>
<td>39.1</td>
<td>5.8</td>
<td>7.1</td>
<td>8.0</td>
</tr>
</tbody>
</table>

Figure 17: Ground truth localization error when apply body shielding impact mitigation

impact) column shows the error estimation in ideal environment generally adopted by existing research work. EvaLoc’s result is obtained with light crowded and body shielding impact.

As we can see from the results, EvaLoc generally provides a better indication of the user’s experience. On the average, the accuracy of EvaLoc differs from the user’s measurement by 17% while the difference between the user’s measurement and the “No Impact” case is 35%.

5.6 Improving Localization Accuracy

So far, the evaluation has focused on EvaLoc’s ability to evaluate localization accuracy. An obvious question is whether it is possible to improve the accuracy by removing some of the environmental effects.

We observe that factors such as body blocking, and to some extent crowd, impact the accuracy when the majority of BSSIDs with stronger RSS are blocked. This is a very challenging problem to address given that many details in the environment are not captured in the fingerprint database.

Inspired by the experience of calibrating a compass by having a user faces different directions to find a bearing field, we observe that if users are also able to capture more fingerprint samples from different directions, it is more likely that BSSIDs with stronger RSSs can be detected.

Note that most of the commercial AP broadcast beacons every 100ms and mobile phones scan through all the 13 WiFi channels to trigger an update to RSS value at roughly every 1.4 seconds if all beacons are detected. In order to mitigate the body blocking effect, the user should at least collect data from two opposite orientations or cover as many orientations as possible. For each of the BSSID collected, if at least one high (unblocked) RSS value can be detected, the accuracy will improve. While it is unlikely that the user can hear from all strong BSSIDs, any additional strong BSSID detected helps.

In the evaluation, for each position, the user collects data from about 4-6 directions. For each BSSID, the strongest RSS is chosen. By picking the strongest RSS for each BSSID in the fingerprint samples, we have a higher chance to remove body or crowd blocking effect. Figure 17 shows the result in the small size lab and large size floor. The average error reduces from 2.45m (GT-Body) to 2.2m (GT-Body-Mitigate) and 6.28m (GT-Body) to 5.1m (GT-Body-Mitigate) for the small size lab and large size floor respectively when the proposed mitigation mechanism is applied.
In conclusion, we see that by merging result from multiple samples in different orientations, we can improve the localization accuracy by limiting the negative impact of self and crowd blocking effect.

6 RELATED WORK
Fingerprint-based Indoor Localization Beginning with the publication of Radar[1], fingerprint-based indoor localization has attracted tremendous efforts from research communities. Early works in indoor localization concentrate on reducing localization error. Horus[20] uses probabilistic techniques to store information about the signal strength distributions and use Bayes inference to estimate the user location with error close to 1 meter. Later works in indoor localization try to reduce the time and effort in constructing fingerprint database or floor plan. Zee[11] exploits WiFi-annotated dead-reckoning and infers location according to the constraint imposed by the floor plan. Walkie-Markie[12] and PiLoc[8] utilize crowd sourced user trajectories to build both floor plan and radio map automatically.

Different with all these works, we want to evaluate the impact of user behavior and environment noise to localization performance. This system has the potential to evaluate the above works in the same configuration.

Localization Performance Evaluation Evaluating the performance of different indoor localization method is attracting more attentions from the research community. EVARilos[15] enables an automated evaluation and comparison of multiple solutions in different environments and using multiple evaluation metrics. [6] systematically analyze the performance achieved by RSSI-based fingerprinting algorithms in different environments to establish a link between the similarities among environments and parametrization of algorithms in these environments. [9] proposed Gaussian process based indoor localization evaluation system, but they are not able to assess localization performance under a more practical scenario with impacts from the user and surrounding environment.

Rather than evaluating the performance of difference system, we focus more on the impact of user behavior and environment noise to fingerprint-based localization performance.

Body Blocking Effect Radar[1] highlights the performance degradation using fingerprint data collected when the user has different orientations. [21] make use of body shielding effect to locate remote AP with commodity mobile phone. The author in [17] suggested to use Least Median of Squares(LMS)[13] estimator to do a robust regression and bound the influence of outlying measurement which contain body shielding impact. However, such robust regression requires a well-calibrated radio map and a relatively stable number of BSSID in each fingerprint. Compared to them, EvaLoc is able to systematically evaluate the impact of body blocking in the scenario of indoor localization.

7 CONCLUSION
In this paper, we study and model the impact of body blocking effect and environment noise to RSS. We then propose EvaLoc, a fingerprint-based indoor localization evaluation system. We use EvaLoc to evaluate the localization performance degradation from body shielding and environment noise. As indoor localization become more and more useful in all kinds of application, we expect such evaluation tools could benefit both researchers and users.

ACKNOWLEDGMENTS
This research is supported by the National Research Foundation, Prime Minister’s Office, Singapore under its International Research Centre in Singapore Funding Initiative.

REFERENCES