

PowerSpy Upgraded: Location Tracking using Mobile Device Power Analysis

Shengtuo Hu, Shibo Chen

University of Michigan

Abstract

Modern mobile devices have many sensors and a variety of information that can be used to infer users' geolocation. In this paper, we reproduce the work of Stanford PowerSpy team on location tracking using power analysis and discuss its availability on newest models. Also, we take different network conditions into consideration. Furthermore, we push PowerSpy a step forward and show that phone's aggregate power consumption from WiFi module is also helpful for location tracking. In the end of this paper, we wrap up lessons we learned from the original paper and our new findings.

1. Introduction

In modern business models, there are many initiatives for companies to track users' location such as advertisement push and price adjustment, etc. Attackers also want to track the victims in pursuit of their evil intention. On the other hand, in concern of safety and privacy, users usually do not want to expose their location information without their consent. Therefore, mobile phone systems consider location information as sensitive and restrict the access to both coarse-grained location from network providers and fine-grained location from GPS. In order to get either coarse-grained or fine-grained location information, an application needs to request the permission and have the user grant the permission manually. Such restriction has greatly limited attackers' ability to conduct attacks. Although the permission may be abused after being granted so that an app can try to trick users to agree its access to location information in the first place, it is out of scope for this paper.

However, victim's motion and location changes may lead to many significant but unexpected effects on mobile phone status, which can be picked up by other sensors (e.g., accelerometer and gyroscope). Das et al. [1] show that these motion sensors can be used to track victim's location. The information exposed by these sensors was originally not considered to correlate with victim's location changes and thus was not well-protected.

A group of researchers in Stanford studied the relationship between mobile device energy consumption rate and victim's location changes [2]. The energy consumption rate on mobile

Email address: {shengtuo, chshibo}@umich.edu (Shengtuo Hu, Shibo Chen)

devices greatly depends on the current signal strength and the hysteresis process. The signal pattern on a specific road is stable and will lead to a similar power consumption pattern (power trace) when the victim is traveling along the road. By collecting power traces when traveling nearby road segments, we can fingerprint a road and then track the victim.

However, such attack has a few constraints. First, the tracking distance needs to be long because short distance travel will not provide enough signal strength changes and thus can not be used to speculate the location. The speed of traveling needs to be high because changes in power consumption would be covered by noises if traveling too slowly. Second, the victim, when being tracked, needs to have a steady phone usage (e.g., navigation, listening to music or idle). A burst of power consumption can be filtered out but a long disruption will disable the attack. Third, when the victim is traveling through the campus, the mobile device may connect to WiFi network which leaves a different power trace, compared to those collected solely via the cellular network.

We also notice that after Android 6.0 upgrade, Doze has been integrated into the platform to optimize power life. Such mechanism prevents background network activity and defers jobs if the user has not interacted with the app for a period. Android 8.0 even restricts background service and broadcast limits for all third-party apps. The app will be suspended, if it doesn't have a foreground activity running, or the phone enters sleeping mode, unless it is on the white list, which may earn the app a few more minutes. These restrictions make attacks more challenging because it is more difficult to record data and transmit the data to the attacker. However, after conducting experiments on the phone running Android 8.0, we argue that such vulnerability is still exploitable if the app is continuously running in the foreground. This kind of attack is still available if the mobile phones are running lower android versions or the attacker can trick the victim to run the app in the foreground when the victim is traveling.

Even though such attacks have been made significantly hard after two major modification on third-party app running policy, we find that such attack is available on devices connecting to WiFi and has cellular network off, which suggests that Android Pads may be vulnerable to such attacks in an indoor environment. Such findings give this attack a new front. In this paper, we first try to reproduce the work that the Stanford team has done and then achieved three other goals:

1. We strengthen their research by providing the proof that network changes, rather than GPS changes, are the major contributor to power consumption variety.
2. We re-evaluate the availability on the new version android device and adjust the threat model.
3. We explore the possibility of tracking the victim in an indoor environment. GPS usually doesn't work in an indoor environment, but indoor environment usually has more environmental varieties that may lead to more dramatic changes in signal strength [3].

2. Threat Models

Popular mobile device systems such as Android or iOS have very restricted permission control on third-party applications accessing to location information. Less sensitive information such as power usage information was not considered to have direct link to location tracking and is often used for apps to optimize battery life or adjust performance. The system may not ask explicitly for user's consent to allow apps to access these types of information. Even the system asks the user for consent, users may grant the permission without a second thought. However, the threat model has changed significantly because of the certain modifications on the system and new findings mentioned in introduction section.

The location of a user can be speculated through power usage if requirements below are met:

1. The victim's activity range is limited and the attacker has pre-knowledge about the victim's activity range so that the attacker can extract fingerprints of all possible routes that the victim may take.
2. During the victim's movement, the victim keeps the app in the foreground. The victim's mobile device has connection to internet and has a stable usage pattern that doesn't introduce significant noise frequently.
3. The victim needs to move a long distance in outdoor. Short distance suffers from a lack of patterns and tracking accuracy drops significantly due to noises. For indoor attacking, the victim's device needs to connect to wifi network and doesn't have cellular network (either turned off or doesn't support).

The attacker can trick the user to install a malware on his or her mobile device and collect power usage information: voltage and current in the background. With such information, the attacker can:

1. Track the victim's location on a specific road in real time or recover this information afterwards in outdoor cellular environment.
2. Speculate which direction the victim takes.

3. Background

In this section, we aim to explain the technical background on the relationship between a phone's location and its cellular power consumption. At first, we describe how location is related to signal strength. Then, we illustrate the relationship between signal strength and power consumption. Further, we explore the major factor to the variety of power consumption. At last, we describe how Hysteresis affect our data collection process.

3.1. Signal Strength and Power Consumption in One Location

Apparently, a phone's signal strength is mainly related to the distance to the base station. According to [4], the signal's power loss is proportional to the square of the distance it travels over. Apart from this path loss caused by the travel distance, the signal strength can also be weakened by objects in the signal path, such as buildings and trees. Finally, reflectors

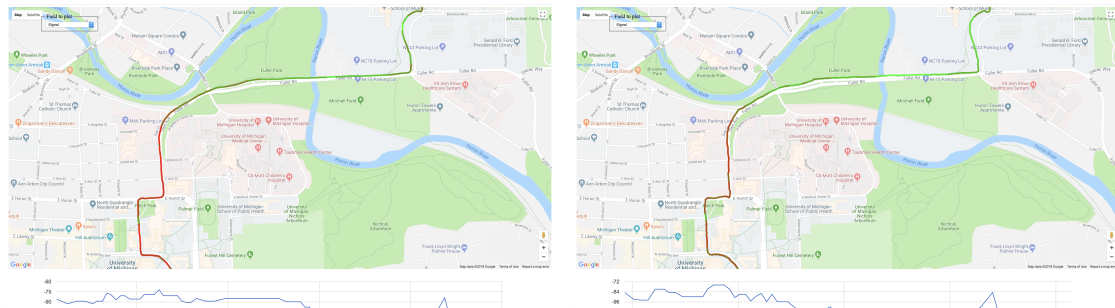


Figure 1: Signal strength traces of the same route measured at different time are stable.

can affect signal strength as well, as these objects cause multi-path interference by reflecting the signal back to the phone through different lengths of paths. Based on these findings, we observe that, in one particular location, signal strength is almost unchanged because base stations, signal obstacles, and reflectors remain stationary.

In addition to the above factors, Schulman et al. [5] show that communication at a poor signal location can lead to the increase of power consumption, compared to a good signal location. The main reason for this circumstance is that the phones power amplifier used for transmission which increases its gain as signal strength drops [4]. Even if a phone is only receiving packets, this effect also occurs, due to the requirement of constant transmission of channel quality and acknowledgments to base stations.

3.2. Stable Signal Strength along a Road

To figure out the connection between the power consumption and the location information, we first demonstrate that signal strength in each location on a road is static over the course of several days. We use an Android app to measure the signal strength along one particular road on different two days. As shown in Figure 1, it is obvious that two signal strength traces are similar with each other.

During these two measurements, we also record the power consumption information, when the phone keeps communicating through the cellular network. Since we have already known that communication at a poor signal location can lead to the increase of power consumption, stable signal strength along a road can result in stable power consumption as well.

3.3. Major Factor of Power Consumption

In the original paper [2], the authors have both cellular network and GPS on when they are conducting experiments, which may be a defect of the paper. They do not provide a solid proof that cellular network is the primary source of the changes shown in the power traces. In this paper, we give such proof by collecting power consumption traces under three conditions: (1) The phone is in an idle state, and all network connections and GPS stay off. (2) The cellular network is on, and location service is entirely turned off. (3) Only location service including GPS is on, and all other networks are turned off by entering airplane mode.

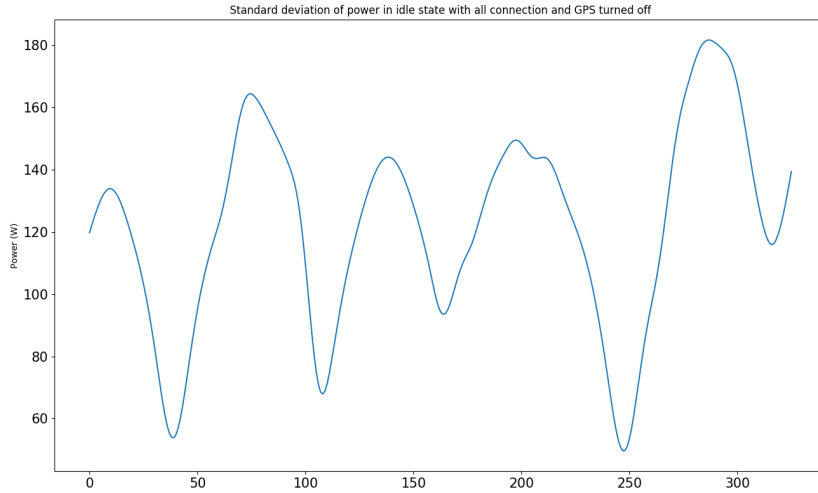


Figure 2: Local standard deviation of power in the idle state with all connections and GPS turned off. Each data point represents the standard deviation of a one-tenth of the all data points in the sample.

We then compare standard deviation of the power traces under these three conditions. Same experiments are also conducted on WiFi network.

The standard deviation of power profile of the device in idle state with all network connections and GPS off is around 130W. Figure 2 shows the trend of local standard deviation of the idle device power profile. Table 1 shows the overall standard deviation of the traces we take. Figure 3 and figure 4 show the trend of local standard deviation of different samples. The GPS-only scenario does not introduce much more varieties than those already existing in the idle state. It is also clear that cellular network and WiFi network typically introduce more changes in the power than GPS does, which network status plays a major role in power consumption changes.

	Route 1 Sample 1	Route 1 Sample 2	Route 2 Sample 1	Route 2 Sample 2
Cellular Only	231.14	253.96	241.47	239.66
GPS Only	164.16	133.13	150.14	140.66

Table 1: Cellular v.s. GPS. Standard deviation of power consumption in terms of major factor, routes and samples in W

	Route 1 Sample 1	Route 1 Sample 2	Route 2 Sample 1	Route 2 Sample 2
WiFi Only	190.39	253.85	290.70	271.84
GPS Only	160.62	104.63	120.61	101.45

Table 2: WiFi v.s. GPS. Standard deviation of power consumption in terms of major factor, routes and samples in W

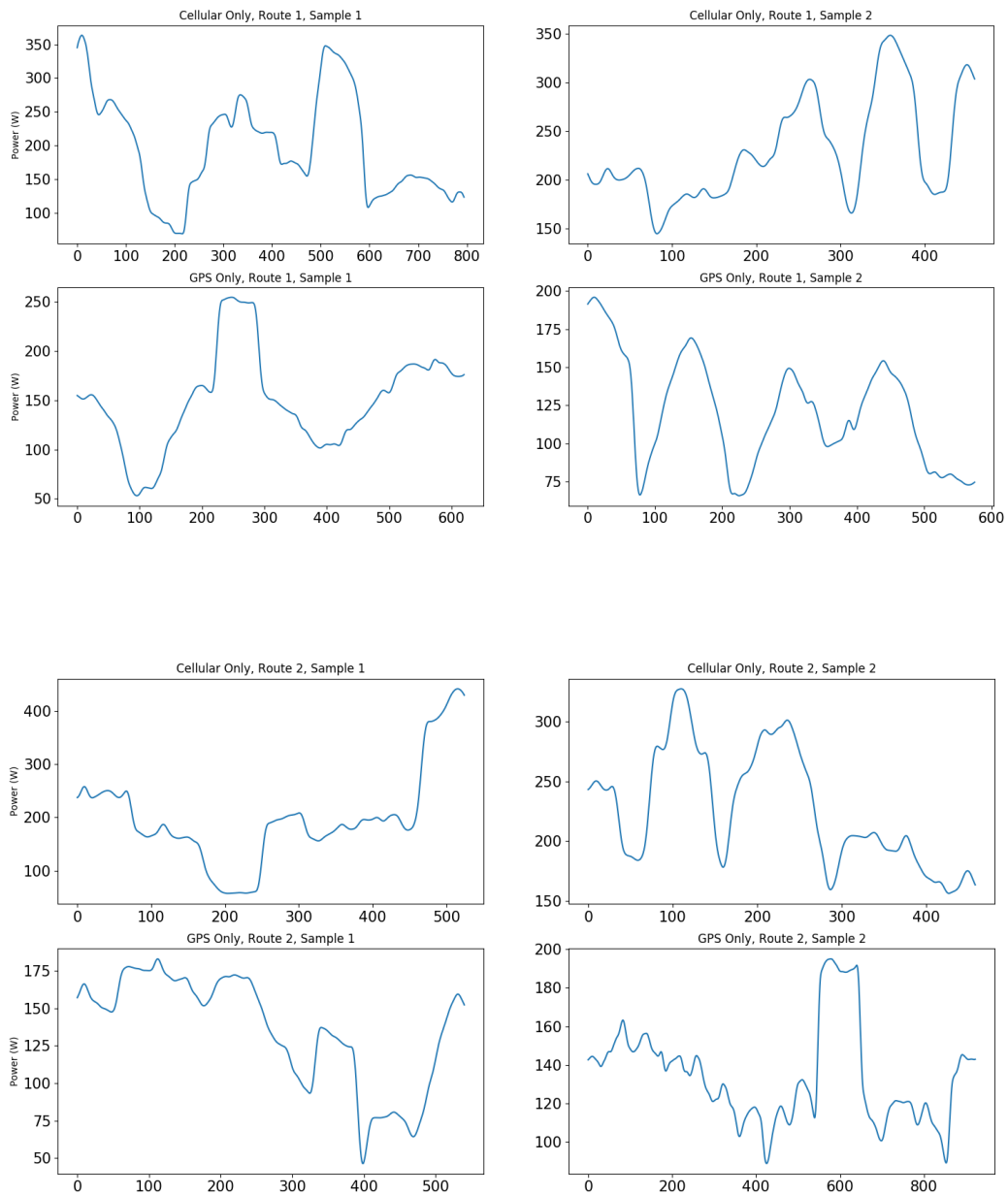


Figure 3: Local standard deviation and the trends of samples between cellular only and GPS only. Each data point represents the standard deviation of a one-tenth of the all data points in the sample.

3.4. Hysteresis

Besides, power consumption information along one road is influenced by the direction of movement as well. On the same road, there are multiple base stations. A phone is connected to the base station having the strongest signal. However, in one location, a phone

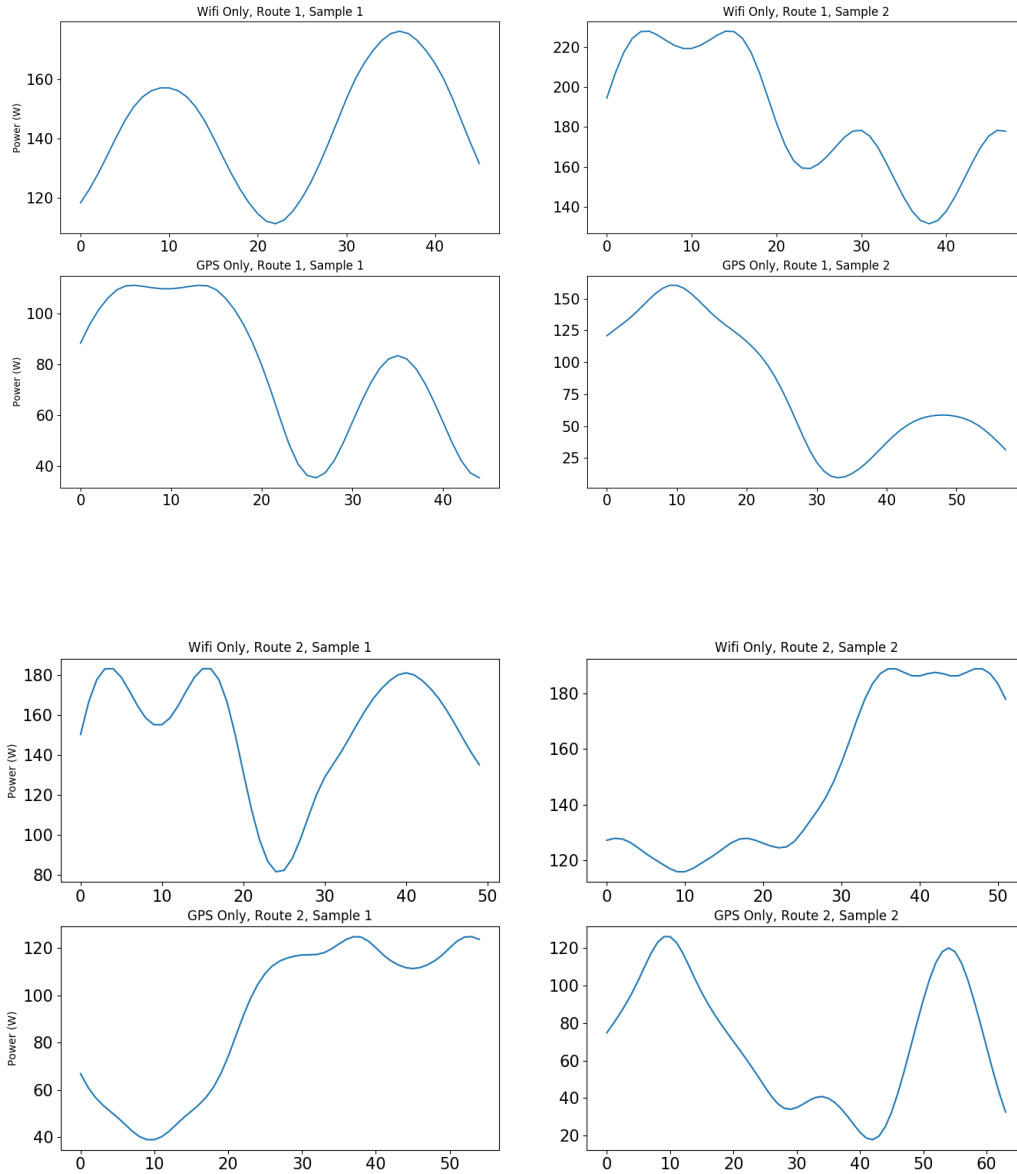


Figure 4: Local standard deviation and the trends of samples between cellular only and GPS only. Each data point represents the standard deviation of a one-tenth of the all data points in the sample.

may connect to different base stations, as the signal strength can be affected by how the phone arrives that location. This is due to the hysteresis algorithm used to decide when to hand-off to a new base station. Hand-off happens only when a phone's received signal strength dips below the signal strength from another base station by more than a given threshold [6]. Therefore, two phones in the same small area can connect to two different base stations. This mechanism means that, for each road, an attacker should collect two

power traces with opposite directions of travel.

4. Methodology

In this section, we describe the methodology to track the location of a user. We first utilize a classification approach to identify the route along which a mobile user is traveling. After that, we conduct the tracking task. We do not assume a particular starting point along the route. The attacker has collected power traces for the target route in advance, and constantly receives new power consumption information from an application installed on the target phone. The attack’s goals include locating the device along the route and continuing tracking it in real-time as it travels along the route.

4.1. Route Distinguishability

In our initial experiments, we have shown that the phone’s power consumption can leak information about the location of a smartphone. We can identify which route the user is taking from a set of pre-collected power traces of different routes. In general, route distinguishability can be described as a classification problem. At first, we collect many power traces and label them with known routes. These pre-collected power traces form the training set. For each power trace, we treat it as a time series. Then, to measure the similarity between two power traces, we assign a score to each comparison of two power traces. Based on these scores, we choose the most likely matching route as the predicted result.

The key component of our methodology is the similarity metric. We have to consider that the pre-collected power traces may vary in both time and length due to two reasons: (1) different rides along the same route can vary in speed at different locations along the ride; (2) routes having the same label can vary slightly at certain points. Therefore, we choose Dynamic Time Warping (DTW) [7] for measuring the similarity between any two power traces. DTW can tolerate misalignment of power traces and can also handle time or speed variants.

For our classification algorithm, we leverage the core idea of k -nearest neighbors (k -NN) algorithm and replace the distance metric with DTW. During the classification phase, the first step is to compute the DTW distance between the testing power trace and all training power traces with known routes (i.e., label or class). After that, we select the known route that has the minimal distance (i.e., $k = 1$) with the testing power trace. The classification problem can be formally described as follows:

$$i = \arg \min_{i \in [1, n]} \text{DTW}(Y, X_i)$$

where X_i is one of the pre-collected power trace, Y is the testing power trace, and n is the total number of training power traces.

Apart from the variants of time and speed of the power trace, we perform normalization to handle issues like different power baselines and variability. Before computing the DTW

distance, we calculate the mean value $mean(X_i)$ and the standard deviation $std(X_i)$ of each trace, and apply the following operation:

$$X_i = \frac{X_i - mean(X_i)}{std(X_i)}$$

where the subtraction and the division are element-wise operations. Furthermore, other pre-processing methods like smoothing and downsampling are applied to reduce noise and computational complexity.

4.2. Real-time Tracking

4.2.1. Tracking via Dynamic Time Warping

In general, this tracking task is similar to the route distinguishability task. The only difference is that we only have part of power trace information while tracking the target device up to this point. Thus, we use the Subsequence DTW algorithm [7] to search a sub-sequence in a larger sequence (i.e., a power trace). In doing so, we can also get the corresponding start and end offsets. For a given tracking task, we search for the sequence of power consumption information we have accumulated since the beginning of the tracking in all the training power traces and select the power trace that has the minimal DTW distance. The predicted location, therefore, is related to the end offset generated by the algorithm. Algorithm 1 presents the logic of the tracking algorithm.

Algorithm 1 Tracking algorithm

```

locked ← false
while target moving do
  loc[i], score ← estimateLocation()
  d ← getDistance(loc[i], loc[i − 1])
  if locked and d > MAX_DISP then
    loc[i] ← loc[i − 1]
  end if
  if score > THRESHOLD then
    locked ← true
  end if
end while

```

4.2.2. Tracking using Optimal Subsequence Bijection

Similar to DTW, Optimal Subsequence Bijection (OSB) [8] can also align two sequences. In DTW, we align two power trace without skipping any elements in the power trace, which means we cannot remove possible noise points. OSB can handle noise points in power traces by skipping some elements.

5. Experiments

5.1. Data Collection

We conduct our experiment at the University of Michigan. The mobile device we use is Moto X4 running Android 8.0 and Google as the carrier. Although Google supports multiple network providers, according to the network profile we extract from the phone, the only network that the experiment device connected to is LTE network provided by T-Mobile.

Route	Cellular Only	Wi-Fi Only	Mixed
<code>bbaits_to_central</code>	6	Not Applicable	3
<code>bbaits_to_north</code>	6	Not Applicable	3
<code>north_route_1</code>	6	6	6
<code>north_route_2</code>	6	6	6
<code>indoor_route_1</code>	6	6	6
<code>indoor_route_2</code>	6	6	6

Table 3: Sample number of each combination

We collect power traces under three environments: (1) on-road taking bus, (2) campus outdoor walking, and (3) campus indoor walking. For each environment, we also consider three types of network condition: (1) cellular only, WiFi turned off, (2) WiFi only, cellular turned off through airplane mode, and (3) mixed network with both mode enabled. For the on-road environment, we only consider network condition (1) and (3) because WiFi turns out to be scarce on the road and also rarely being connected to when driving through at high speed. For each combination, we collect at least two routes to assess distinguishability and multiple traces to reduce the effect of unexpected noises. The traces we profiled and their corresponding names are listed in Figure 5.

The speed on `bbaits` bus is uncontrollable, which has an impact on mixed network environment because the mobile device will not connect to nearby WiFi when the bus drives through fast but is more likely to connect to WiFi if we pass the signal source in a slow speed. However, we argue that such variant actually makes the experiment more practical because the victim will have different movement speed and should be able to be overcome by OSB.

Since the WiFi signal is a near signal source comparing to the cellular signal, a few meters difference can cause a significant difference in signal strength and WiFi connection point. A few meters off when collecting the data is inevitable, and its effect on location tracking is to be determined.

GPS is extremely inaccurate in indoor environments so that we are currently unable to get the correlation between location and energy. Therefore, for indoor environments, we only evaluate route distinguishability.

5.2. Route Distinguishability

To evaluate the classification task for route distinguishability, as indicated in Table 3, we collect several power traces for two different routes under different network settings. In total,

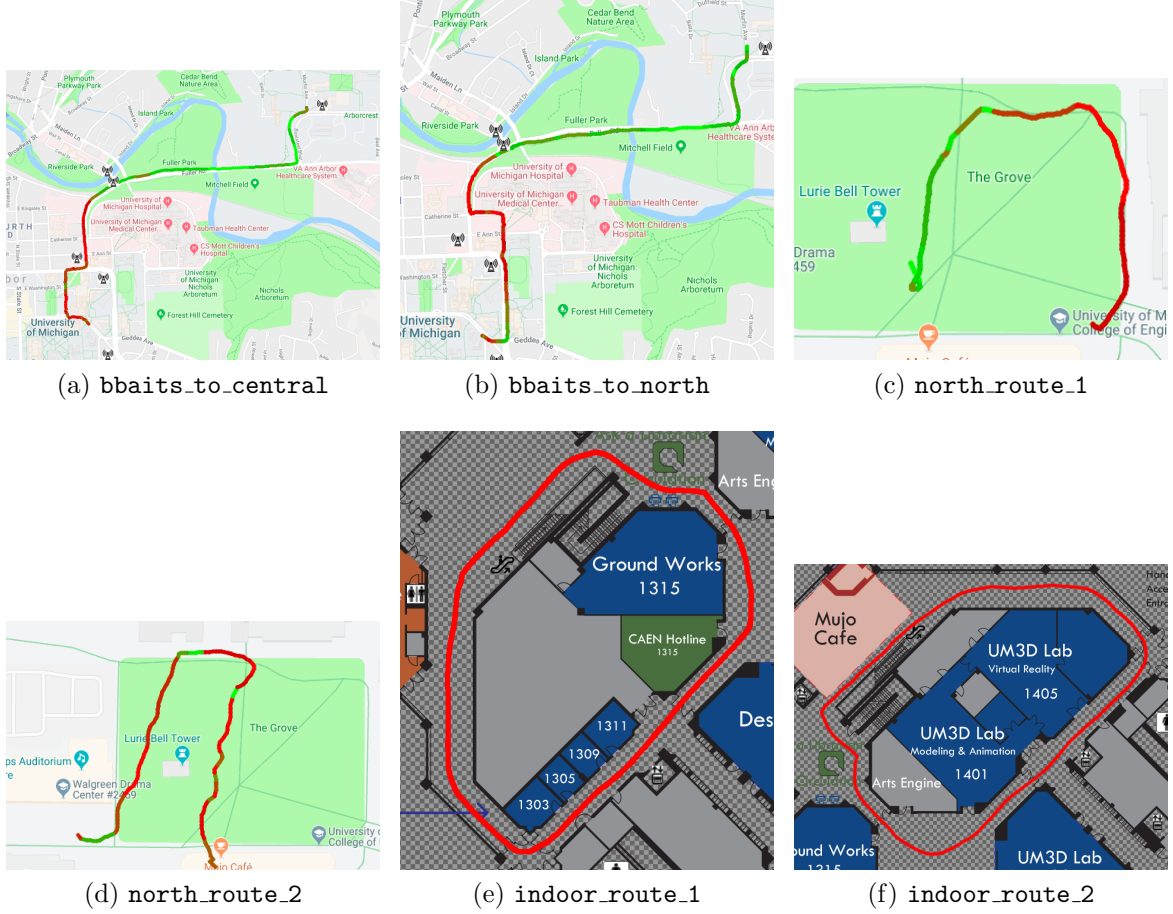


Figure 5: Routes and their corresponding names

we have three sets of experiments for route distinguishability: (1) outdoor + bus (Figure 6), (2) outdoor + walk (Figure 7), and (3) indoor + walk (Figure 8). Further, we consider network settings, as presented in subfigures of each set.

As shown in Figure 6, the classifiers can identify almost all the testing traces correctly. The results are consistent with the original PowerSpy paper.

In Figure 7, we find that all three experiments show bad classification accuracy. For the cellular network, the moving area is quite small. We dig into the raw data and notice that, during each data collection phase, the mobile phone is attached to the same base station. Thus, there does not exist obvious data changes. For WiFi network,

In Figure 8, we find that the route distinguishability is not so good if the victim is under either cellular network or mixed network, but it is good under WiFi-only condition. This circumstance may be due to the stable change pattern of WiFi access points. As well, enough number of WiFi access points along the route may provide sufficient data for route distinguishability task. Besides, cellular signal strength at indoor environment is quite bad, which can lead to high power consumption. Therefore, for cellular network condition, the mobile

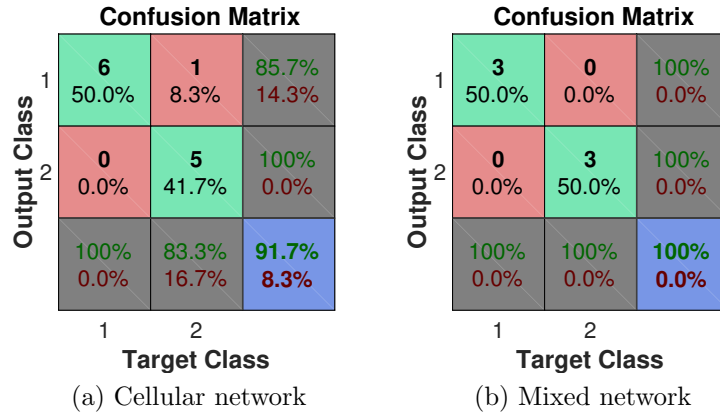


Figure 6: Outdoor + Bus. Class 1 stands for bbaites_to_central, and class 2 for bbaites_to_north.

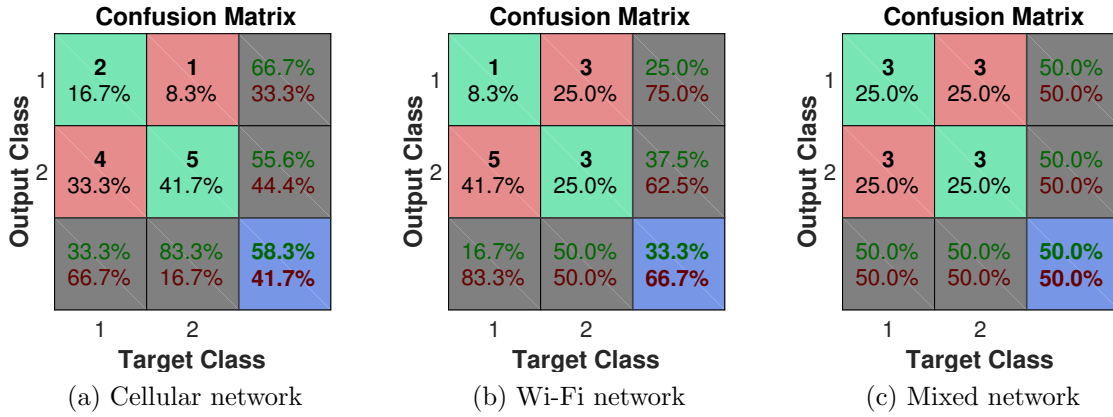


Figure 7: Outdoor + Walk. Class 1 stands for north_route_1, and class 2 for north_route_2.

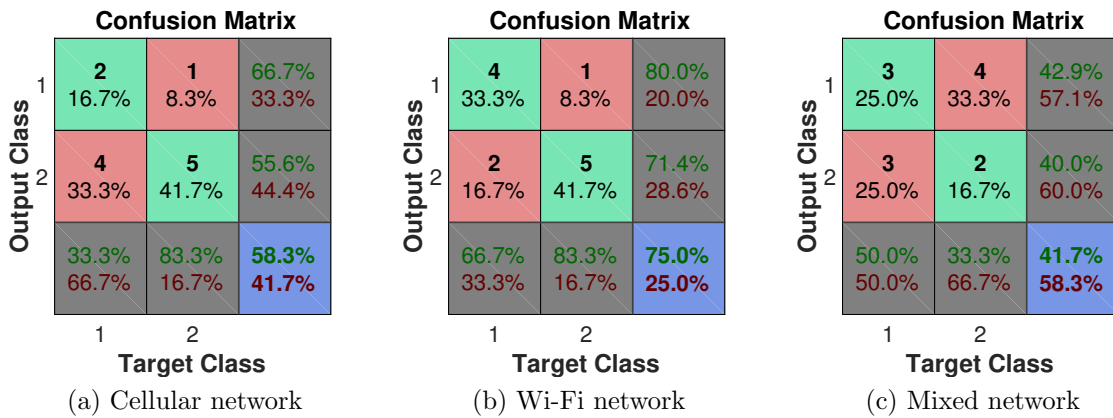


Figure 8: Indoor + Walk. Class 1 stands for indoor_route_1, and class 2 for indoor_route_2.

device cannot even communicate through the network, and the high power consumption may introduce noise data points. Also, for mixed network condition, the power consumption of cellular modem may overwhelm the power information of WiFi module, so the classifier cannot extract enough meaningful information for route distinguishability task. Such finding suggests that this attack also works indoor. Even though we are unable to conduct real-time tracking for indoor movement, we argue that since indoor area is small comparing to outdoor area, knowing the route the victim has taken would reveal much information about where the victim might be.

5.3. Real-Time Tracking

We also evaluate Algorithm 1 for real-time mobile device tracking using a set of 5 training profiles and an additional test profile. Figure 9 shows that the estimated error is small compared to the original paper, which claims that 90% of estimated errors are less than 1km. For more sets of experiments of real-time tracking, please refer to the Appendix Appendix A.

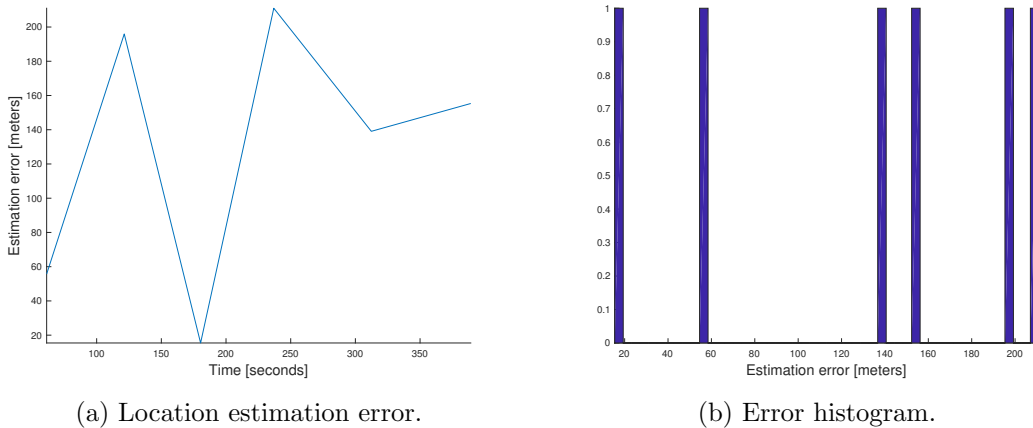


Figure 9: Estimation errors for motion-model tracking.

6. Discussion

6.1. Strength and Weakness Comparing to the Original Work

Based on the original PowerSpy paper, we have three improvements:

1. Both PowerSpy team and we have both network and GPS on when collecting training data set and testing data set because we need both power consumption status and location to make reference and comparisons. But the victim may not have the GPS service on. While PowerSpy team did not provide any proof to support the argument that it is network condition, rather than GPS, causes the difference, we analyze the effect that network condition and GPS have on power traces separately and confirmed that GPS had negligible effects comparing to the effects that network condition have on the power profile.

2. While PowerSpy team take power traces for a long distance continuously and then cut it into different routes, which may have an unknown effect the power traces collected. We take power traces one by one separately, which should be more convincing.
3. We provide experiments based upon different combination of geo-environments and network conditions. We get a positive result on distinguishing routes indoor under WiFi network, which extends the old attack model.

However, we also have two drawbacks comparing to the original paper:

1. Due to the lack of devices to test on, we only conduct experiments on one phone. The original work has data traces taken from different devices of the same model and different models to show that power traces stay relatively stable on the same route.
2. The size of data set is relatively small comparing to the original work because of realistic constraints on mobility and the number of devices. We have only two routes and a short travel distance.

6.2. Limitations

In addition to the limitations discussed above, we have such limitations:

1. Different carriers would have different configurations and arrangements of their base stations. Based on our observation, it is actually subject to change so that such attack requires the attacker to collect the most recent power traces using the same carrier as the victim in order to conduct this attack.
2. Since GPS signal is actually inaccurate indoor, we are unable to use GPS as location references so that we are unable to exploit trackability indoor.
3. Even though we have proved that GPS has a relatively small effect on power consumption, we are unable to rule out the effect of GPS.

7. Related Works

Power analysis is a form of side channel attack, which is originally applied in cryptography. Kocher et al. [9] introduced Simple Power Analysis (SPA) and Differential Power Analysis (DPA) to the open cryptology community. They extracted high sample rate (20 MHz) power traces from externally connected power monitors to recover correct private encryption keys from a cryptographic system. Researchers [10, 5] also uncovered the relationship between signal strength and power consumption in smart phones. Bartendr [5] further showed that signal strength along a path is stable across several drives.

PowerSpy [2] combines these insights together and successfully extracts users' location information. The original authors monitor the cellular modem's changes in power consumption with a unprotected internal power monitor, whose sample rate (100 Hz) is much lower than the requirement of SPA or DPA. In our work, we push PowerSpy a step forward and extend the attack scenario. Our newer version of PowerSpy introduces WiFi module as another source of power consumption so that we can track users' location indoor, especially when users are playing tablets without cellular modems.

Prior works have demonstrated that cellular modems and WiFi module can leak location. However, all of these methods [11, 12, 13, 14] require the permission of accessing signal strength and base station ID or WiFi network name (SSID). PowerSpy does not need this access permission and only relies on unprotected power consumption information.

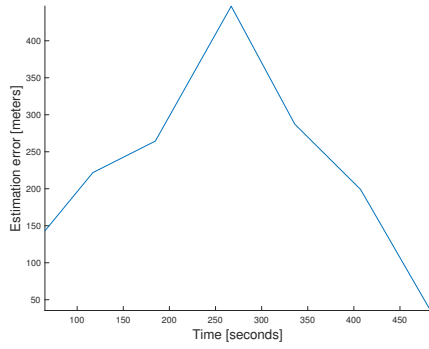
8. Conclusion

Three years after the publication of the original work, it becomes more challenging to conduct such attacks due to new constraints imposed on the third-part apps and thus, the threat model has changed significantly. However, information is still leaked out during the reproduction, which implies no defense has been deployed on either hardware level or system level.

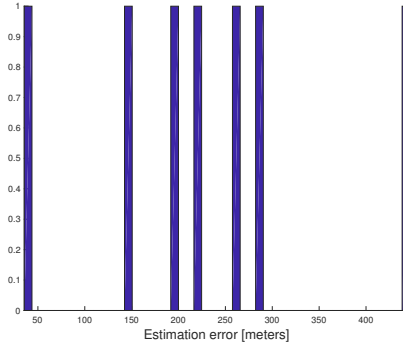
Furthermore, we find that such attack is also available under indoor and WiFi-only condition. Such finding does not only extend the threat model but also draws attention to what else may be leaked through power consumption information.

- [1] N. B. . M. C. Anupam Das, Tracking mobile web users through motion sensors: Attacks and defenses, 2016.
- [2] G. N. G. A. V. . D. B. Yan Michalevsky, Aaron Schulman, Powerspy: Location tracking using mobile device power analysis, 2015.
- [3] S. S. A. M. R. R. C. A. E. M. F. Wang, He, M. Youssef, Unsupervised indoor localization, 2012.
- [4] A. Goldsmith, Wireless Communications, Cambridge university press, 2005.
- [5] A. Schulman, V. Navda, R. Ramjee, N. Spring, P. Deshpande, C. Grunewald, K. Jain, V. N. Padmanabhan, Bartendr: a practical approach to energy-aware cellular data scheduling, in: Proc. of MOBICOM, 2010.
- [6] G. P. Pollini, Trends in handover design, 1996.
- [7] M. Müller, Information retrieval for music and motion, Springer, 2007.
- [8] L. J. Latecki, Q. Wang, S. Koknar-Tezel, V. Megalooikonomou, Optimal subsequence bijection, in: Proc. of ICDM, 2007.
- [9] P. Kocher, J. Jaffe, B. Jun, Differential power analysis, in: Proc. of CRYPTO, 1999.
- [10] A. Carroll, G. Heiser, An analysis of power consumption in a smartphone, in: Proc. of USENIX ATC, 2010.
- [11] J. Krumm, E. Horvitz, LOCADIO: inferring motion and location from wi-fi signal strengths, in: Proc. of MobiQuitous, 2004.
- [12] K. Muthukrishnan, B. van der Zwaag, P. J. M. Havinga, Inferring motion and location using WLAN RSSI, in: Proc. of MELT, 2009.
- [13] W. R. Ouyang, A. K. Wong, C. A. Lea, V. Y. Zhang, Received signal strength-based wireless localization via semidefinite programming, in: Proc. of GLOBECOM, 2009.
- [14] T. Sohn, A. Varshavsky, A. LaMarca, M. Y. Chen, T. Choudhury, I. E. Smith, S. Consolvo, J. Hightower, W. G. Griswold, E. de Lara, Mobility detection using everyday GSM traces, in: Proc. of UbiComp, 2006.

Appendix A. Real-Time Tracking Results

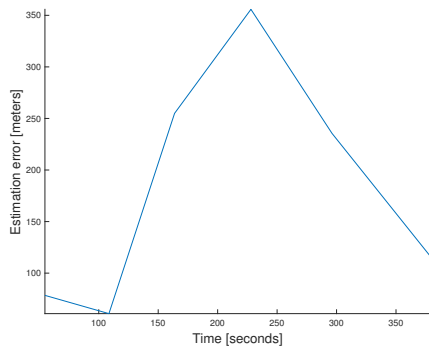


(a) Location estimation error.

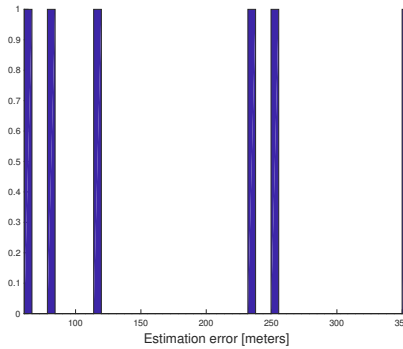


(b) Error histogram.

Figure A.10: Estimation errors for motion-model tracking.

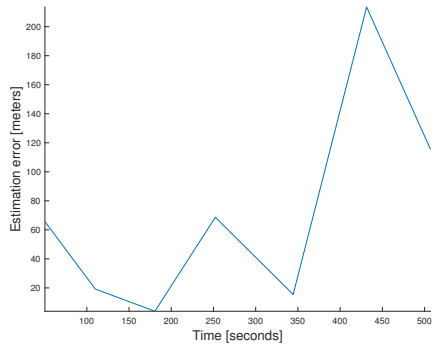


(a) Location estimation error.

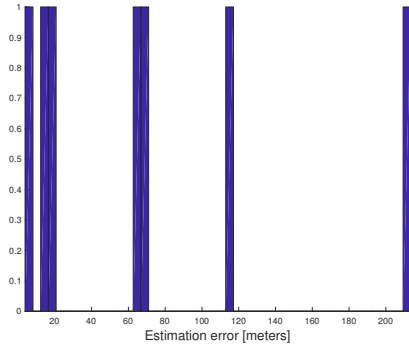


(b) Error histogram.

Figure A.11: Estimation errors for motion-model tracking.

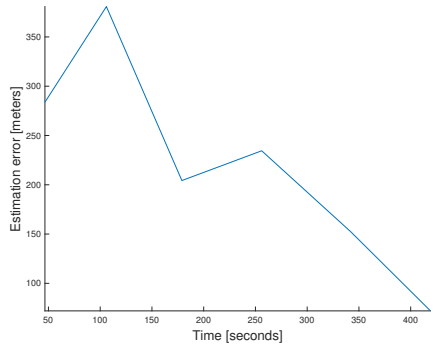


(a) Location estimation error.

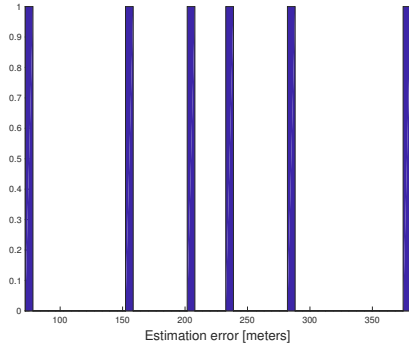


(b) Error histogram.

Figure A.12: Estimation errors for motion-model tracking.

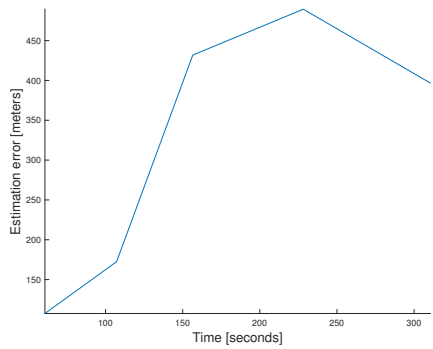


(a) Location estimation error.

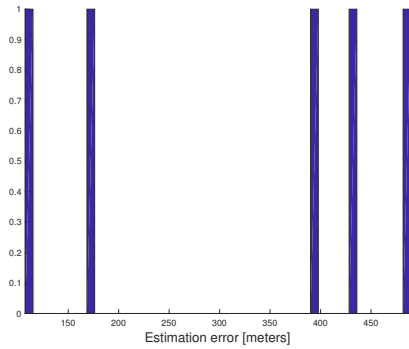


(b) Error histogram.

Figure A.13: Estimation errors for motion-model tracking.



(a) Location estimation error.



(b) Error histogram.

Figure A.14: Estimation errors for motion-model tracking.