Agility Maneuvers to Mitigate Inference Attacks on Sensed Location Data

Giuseppe Petraccagxp18@cse.psu.edu
Computer Science and Engineering
The Pennsylvania State University

Lisa M. Marvelmarvel@ieee.org
Ananthram Swamiananthram.swami.civ@mail.mil
Army Research Laboratory

Trent Jaeger	tjaeger@cse.psu.edu
Computer Science and Engineering
The Pennsylvania State University

Abstract—Sensed location data is subject to inference attacks by cybercriminals that aim to obtain the exact position of sensitive locations, such as the victim’s home and work locations, to launch a variety of different attacks. Various Location-Privacy Preserving Mechanisms (LPPMs) exist to reduce the probability of success of inference attacks on location data. However, such mechanisms have been shown to be less effective when the adversary is informed of the protection mechanism adopted, also known as white-box attacks. We propose a novel approach that makes use of targeted agility maneuvers as a more robust defense against white-box attacks. Agility maneuvers are systematically activated in response to specific system events to rapidly and continuously control the rate of change in system configurations and increase diversity in the space of readings, which would decrease the probability of success of inference attacks by an adversary. Experimental results, performed on a real data set, show that the adoption of agility maneuvers reduces the probability of success of white-box attacks to 2.68% on average, compared to 56.92% when using state-of-the-art LPPMs.

I. INTRODUCTION

Researchers have long discussed privacy concerns rising from inference attacks on location data [6], [7], [8]. An inference attack is a data mining technique performed by analyzing data to illegitimately gain knowledge about a subject. For instance, an adversary can infer sensitive locations (i.e., victim’s work or home) by monitoring the data points produced, within a time period, by the victim’s mobile platform through GPS or Wi-Fi signals [2], [3]. An example is presented in Figure 1 where sensitive locations are marked by a dashed perimeter, whereas location data points available to the adversary are presented as gray circles. Upon obtaining the victim’s home and work locations, an adversary could physically harm the victim or violate the victim’s privacy in several ways.

Researchers have proposed various Location-Privacy Preserving Mechanisms (LPPMs) such as Spatial Cloaking [2], which removes data points that are inside a circular region around a point marked as sensitive by the mobile platform owner. Furthermore, noise, such as Gaussian [2] or Laplacian, can be added to generate noisy data points. Distortion can be used to add random noise to location data and avoid releasing actual locations. Reduced Sampling can reduce the sampling interval to decrease the amount of collected location data. Lastly, Rounding [2] can be used to round data point values and reduce accuracy. An example of the effects of applying such protection mechanisms to location data is reported in Figure 2. To an extent, these techniques have been shown to be able to reduce the probability of success in identifying sensitive locations when the adversary does not know about the protection mechanisms adopted [2], [3], also known as black-box attacks. However, such mechanisms might be less effective when the adversary is informed of the adopted mechanisms to protect sensitive locations, also known as white-box attacks.

In this paper, we first analyze how well existing LPPMs can protect sensitive locations against white-box attacks. We then propose and evaluate the effectiveness of three new protection mechanisms based on the use of agility maneuvers (e.g., alteration of the environment in response to adversarial action and perceived threat [11]) to better address white-box attacks.

II. ADVERSARY MODEL

In this section, we specify the information available to the adversary while performing inference attacks on location data. In both black-box and white-box attacks, the adversary has access to location data points (time-stamped latitude and longitude coordinates) produced by GPS and Wi-Fi receivers on the victim’s mobile platform (e.g., smartphone). In addition, in white-box attacks we assume a powerful adversary who knows not only the mechanism used to protect location data, but also the configuration of the protection mechanism (e.g., parameters used as input to the protection mechanism). For example, if spatial cloaking has been used, the adversary would know the radius used to define the circular region around locations marked as sensitive (a complete list is provided in Table I). Furthermore, we assume the adversary can adopt four heuristics, also used in related work [2], to perform
inference attacks. The first heuristic, named First and Last Destination, assumes the victim would probably go to work as first destination in the morning and go home as the last destination at night, therefore, the adversary identifies where the victim moves to at the beginning of the day, and where the victim usually terminates the daily journey. The second heuristic, named Most Stationary Way Points, assumes the victim spends more time at home and at work than at any other location, therefore, the adversary identifies the most stationary data points by calculating the amount of time spent in a fixed location until the next data point is recorded. The third heuristic, named Larger Clusters, assumes that most of the victim’s data points will be around home and at work respectively, therefore, the adversary identifies the two clusters with the largest amount of data points. Finally, the fourth heuristic, named Best Time, assumes the victim stays at work and sleeps at home during specific time intervals, therefore, the adversary isolates data points during those time intervals.

We assume that anonymization of data is used to protect the identity of the mobile platform owner and exclude simple attacks where the adversary uses web search engines or similar approaches to identify the victim’s home or work address.

III. ADOPTION OF AGILITY MANEUVERS

We now introduce the use of agility maneuvers activated upon the occurrence of environmental events to mitigate the probability of success of black-box and white-box attacks using location data to infer the victim’s home and work locations.

A. Agility Maneuvers Activation

In system security, agility maneuvers are systematically activated in response to specific system events (i.e., internal state of sensors) to rapidly and continuously control the rate of change in system configurations and increase diversity in the space of readings, which would decrease the probability of success of inference attacks by an adversary. In our experimental study, we propose and investigate the activation of agility maneuvers when: (1) the sensed data becomes stationary, therefore there is no substantial movement; and (2) no new sensed data points are produced within a short time interval. We have chosen these events for the activation of agility maneuvers for the following two reasons. First, when the data points become stationary, they start leaking more information regarding places where the victim spends more time, which includes work and home locations. Second, when there is no new sensed data, there is a chance to introduce synthetic\(^1\) data points that potentially would not impact legitimate use of such data but would rather mislead an adversary constantly monitoring the victim’s movements. The investigation of the impact on legitimate use and the study on where to place such location security solutions, either at the device or at the edge location server [10], [11], is part of future work.

B. Agility Maneuvers Selection

In this subsection, we present the three agility maneuvers proposed as defenses against white-box attacks.

The first maneuver is Random Obfuscation, which focuses on the rate of change for system configurations. This maneuver randomly selects one protection mechanism, from the set of available mechanisms (e.g., spatial cloaking, noise, distortion, rounding and reduced sample rate), every time the sensing data becomes stationary for a prolonged time period. A snippet of the Random Obfuscation Algorithm, with six different protection mechanisms, is reported in Algorithm 1.

---

\(^1\)Fake data points opportunistically crafted.
be difficult to infer sensitive locations among a uniform distribution of data points as explained in Section III-C, and syn_lon and syn_lat represent the synthetic data points provided as output besides real data points.

Algorithm 3 Temporal Distribution

Require: min_lat, max_lat, min_lon, max_lon and time_threshold
1: time_threshold = n (seconds) \small Time Interval
2: while LocationDataCollectionEnabled do
3:     if (current_time – last_data_time) ≥ time_threshold then
4:         syn_lat ≜ Distribute(lat, min_lat, max_lat)
5:         syn_lon ≜ Distribute(lon, min_lon, max_lon)
6:         Output(syn_lat, syn_lon)
7:     else
8:         if NoNewDataPoints do
9:             syn_lat ≜ Distribute(lat, min_lat, max_lat)
10:                syn_lon ≜ Distribute(lon, min_lon, max_lon)
11:                 Output(syn_lat, syn_lon)
12:             else
13:                 Output(lat, lon)
14:         end if
15:     end if
16: end while

Generating synthetic data points to uniformly distribute data points around the space of readings, whenever there is no new real data within a short time interval, would avoid time-stamped information from revealing useful information to an adversary. Figure 5 depicts how the set of data points (different shades for different time frames) changes when using the third agility maneuver compared with the original set of data points reported in Figure 1.

The third maneuver is Temporal Distribution, which focuses on deception of the adversary constantly reading the victim’s location. It consists of uniformly redistributing data points in the space of readings by generating synthetic data points whenever the location sensors (GPS and Wi-Fi receivers) do not produce new data within a short time interval due to out of reach location or lost signal. A snippet of the Temporal Distribution Algorithm is reported in Algorithm 3 where the Distribute function implements a uniform distribution of data points as explained in Section III-C, and syn_lon and syn_lat represent the synthetic data points provided as output besides real data points.

---

Fig. 3: Location data points around Home (left) and Work (right) with Random Obfuscation.

Fig. 4: Location data points around Home (left) and Work (right) with Spatial Distribution.

Fig. 5: Location data points around Home (left) and Work (right) with Temporal Distribution. Different shades represent different time periods (one hour granularity).

3Provided as output besides real data points to achieve diversity and uniform distribution in the space of readings.

3New or modified data points provided as output besides real data to deceive an adversary.
C. Distribution of Data Points

We now describe the distribution of synthetic data points among the real data points to implement the agility maneuvers presented in the previous section.

Before giving more details about the distribution of data points, we motivate the use of such distribution with the following observation. Inference attacks require high accuracy and continuous reading of data for a prolonged time period in order to allow an adversary to reconstruct sensitive information, so a uniform distribution of data points along the entire space of readings would increase the number of guesses an adversary has to make even for white-box attacks. In fact, a uniform distribution of data points in the entire space of readings would hide spatial and temporal patterns otherwise visible to an adversary. By using uniform distribution, all data points within a selected geographical area of interest have the same probability of appearing in the data set available to the adversary, as shown by the equation and plot of the probability distribution function $f(lat)$ below:

$$f(lat) = \begin{cases} \frac{1}{max_lat-min_lat} & \text{for } min_lat \leq lat \leq max_lat \\ 0 & \text{for } lat < min_lat \text{ or } lat > max_lat \end{cases}$$

The two parameters $(min_lat$ and $max_lat$) are respectively minimum and maximum latitude (similarly we have $min_lon$ and $max_lon$ for longitude) selected in order to achieve uniform distribution within a specific geographic area.

Simply applying uniform distribution to data points over the entire space of readings would create unrealistic synthetic data, identifiable by a more advanced adversary. To achieve a realistic distribution of data points we propose an incremental uniform distribution that subdivides the entire set of readings into zones. As shown in Figure 6, the entire space of readings (e.g., latitude and longitude on a 2D map) is divided in zones. Starting from the zone containing the current real data (e.g., zone $(x,y)$), we systematically add synthetic data to achieve uniform distribution within the zone. Once a number of data points, equal to a set threshold $\tau$, is reached within the current zone then an adjacent zone (i.e., zone $(x-1, y)$) is selected to extend the current zone and increase the area covered by the uniform distribution. This approach gradually achieves an incremental uniform distribution that better simulates real movements while placing synthetic data among real data.

![Incremental Uniform Distribution](image)

**Fig. 6: Incremental Uniform Distribution. Upon a number of data points, equal to a set threshold $\tau$, is reached in zone $(x,y)$ then the current zone is extended by adding an adjacent zone, for instance zone $(x-1, y)$ in the reported example.**

<table>
<thead>
<tr>
<th>Protection Mechanism</th>
<th>Information Known by Adversary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial Cloaking</td>
<td>Radius of Circular Region</td>
</tr>
<tr>
<td>Gaussian Noise</td>
<td>Mean Value and Standard Deviation</td>
</tr>
<tr>
<td>Laplacian Noise</td>
<td>Mean Value and Standard Deviation</td>
</tr>
<tr>
<td>Distortion</td>
<td>Exact Digits affected by Distortion</td>
</tr>
<tr>
<td>Reduced Sampling</td>
<td>Sampling interval (Rate)</td>
</tr>
<tr>
<td>Rounding</td>
<td>Number of Truncated (Less Significant) Digits</td>
</tr>
<tr>
<td>Random OffLocation Spatial Distribution</td>
<td>Zone Size and Threshold, Min/Max Latitude and Longitude</td>
</tr>
<tr>
<td>Temporal Distribution</td>
<td>Zone Size and Threshold, Min/Max Latitude and Longitude</td>
</tr>
</tbody>
</table>

**TABLE I: Information available to the adversary in white-box attacks.**

IV. EXPERIMENTAL EVALUATION

A. Data Set Description

All of the experiments described in this paper were performed using our CampusLife data set, a collection of over 483,840 time-stamped location data points collected by using GPS and Wi-Fi signals around the Penn State University Campus at University Park, PA. We collected location data by using the GPSLogger app [9] on a Nexus 5X smartphone running Android 6.0.1. The data collection lasted 4 weeks 24 hours/day. The data set reports location data relative to all movements performed by a graduate student working on campus and living off campus, and it is divided in daily reports where each time-stamped data point has the following format:

<date,time,latitude,longitude,provider>

The provider field is either gps or network, based on the signal used (GPS or Wi-Fi) to derive the subject’s position.

B. Protection Mechanisms and Adversarial Action Modeling

We implemented six location-privacy preserving mechanisms (described in Section I and Figure 2), by following the description reported in previous related work [2], and three new algorithms for the proposed agility maneuvers (Section III-B). Lastly, we implemented the four heuristics (presented in Section II) to simulate the adversary’s action in black-box and white-box attacks. Additionally, in white-box attacks, we used reconstruction functions to simulate the adversary’s ability of reconstructing an approximation of the original data points when protection mechanisms are used. For each protection mechanism, we designed a reconstruction function that gets as input the information available to the adversary, for that specific protection mechanism (e.g., radius for Spatial Cloaking or mean and standard deviation for uniform distribution), and returns as output an approximation of the original data points. A summary of the complete information available to the adversary, based on the adopted protection mechanism, is reported in Table I. All algorithms were implemented in Python and the source code is made available.

C. Experimental Setup

We started our experiments by analyzing the effects of the six location-privacy protection mechanisms (highlighted in

4We used our CampusLife data set because other publicly available data sets do not report information about subjects’ home and work location for anonymization and privacy reasons, however this information represents the ground truth required to validate experimental results. The data set is available for download at [http://sites.psu.edu/petracca/campuslife/](http://sites.psu.edu/petracca/campuslife/).

5[http://sites.psu.edu/petracca/location_priva...](http://sites.psu.edu/petracca/location_privacy_code/)
Section I) when applied to the original data points available in our CampusLife data set. Each defense mechanism had as input the original data set and produced a modified set of readings to implement the appropriate defense protection mechanism. We tested the four heuristics (discussed in Section II) given each of the modified set of readings to simulate black-box attacks. We then reconstructed an approximation of the original data points and tested the same four heuristics given the new data points, to simulate white-box attacks. Table I summarizes the information available to the adversary in white-box attacks. We finally tested the same four heuristics given the original set of readings from our CampusLife data set modified to implement each of the three proposed agility maneuvers (for black-box attacks). We then reconstructed an approximation of the original data points and tested the same four heuristics given the new data points (for white-box attacks). We measured the percentage of success as the number of times an attacker succeeded in identifying both victim’s home and work locations over 28 days (4 weeks) worth of data. The attacker used a daily report of location data to make a single guess (per day) of the victim’s home and work locations.

D. Experimental Results for LPPMs

Clearly, the use of LPPMs decreases the number of times the adversary is able to infer the victim’s home and work locations by reducing the percentage of success for black-box attacks down to 3.57% (best case for Rounding) and on average 39.57% (results are summarized in rows 2-7 in Table II). However, these mechanisms\(^6\) are considerably less effective against white-box attacks, with a measured average percentage of attack success up to 56.92%.

In particular, Spatial Cloaking performs best against black-box attacks that use the Best Time heuristic (3.57%). However, it performs considerably worse against white-box attacks (on average 63.39%) because the adversary can identify the area around sensitive location (by means of geometry calculations) by knowing the radius of the hidden area around the sensitive location. Gaussian Noise and Laplacian Noise perform best against black-box attacks using Most Stationary Way Points (on average 55.35% Gaussian and 49.99% Laplacian) or Larger Clusters (on average 53.57% Gaussian and 41.06% Laplacian) heuristics, with Laplacian Noise being slightly better (on average -8.95%) because the Laplace distribution has heavier tails than the Gaussian distribution. However, their performance considerably degrades (on average 79.9% Gaussian and 75.89% Laplacian) for white-box since an adversary can cancel out the noise applied to the original data points. Distortion performs better than Gaussian and Laplacian Noise in both black-box (on average 29.02%) and white-box (on average 30.35%) attacks. It also performs better (on average -33.04%) than Spatial Cloaking in white-box attacks. Furthermore, Distortion is not heavily affected (on average only 1.33% more for white-box respect to black-box attacks) by the amount of information available to the adversary. This is due to the randomness used to cause distortion of real data points, which is not totally reversible. Reduced Sampling is one the least effective mechanisms with an average percentage of success of 82.59% for both black and white-box attacks. This is because reducing the number of data points available to the adversary is not sufficient to hide specific patterns. Finally, Rounding is slightly worse (on average +1.34%) than Spatial Cloaking for black-box attacks, but much better (on average -55.8%) than Spatial Cloaking for white-box attacks. This is because, the information lost by rounding the data points cannot be reconstructed by the adversary with accuracy.

E. Experimental Results for Agility Maneuvers

Agility maneuvers are less affected by the amount of information available to the adversary. In fact, agility maneuvers are effective in reducing the percentage of success of both black-box attacks (overall average\(^7\) 13.13%) and white-box attacks (overall average\(^8\) 13.72%) on location data (results are summarized in rows 8-11 in Table II).

In particular, for white-box attacks Random Obfuscation performs better (-14.04%) than most of the previously analyzed protection mechanisms on average. This is mainly due to the randomness used in selecting the protection mechanisms activated during a specific time frame among those available. However, it is slightly less effective against white-box attacks (on average 42.40%) compared to black-box attacks (on average 40.05%). Further, it is much less effective than the other two agility maneuvers. On average, the percentage of success of black-box attacks increases 35.59% more than Spatial Distribution and 34.37% more than Temporal Distribution. For white-box attacks, the percentage of success increases 37.94% more than Spatial Distribution and 37.05% more than Temporal Distributions. Spatial Distribution, on average, performs slightly better (-0.98%) than Temporal Distribution, with 4.46% average percentage of success for both black-box and white-box attacks. Temporal Distribution, however, has an average percentage of success of 5.35% for both black-box and white-box attacks. In particular, Spatial Distribution performs best (only 3.75%) against Most Stationary and Larger Clusters heuristics. Temporal Distribution performs best (only 3.75%) against First/Last Destination and Best Time heuristics. Interestingly, both Spatial and Temporal Distribution remain stable even in white-box attacks because a uniform distribution of data points over the set of readings increases the number of possible values an adversary have to chose from.

Our experimental results confirm that Spatial and Temporal Distribution outperform other protection mechanisms. In fact, on average, they perform better (-34.67% for black-box and -52.02% for white-box attacks) than other location-privacy protection mechanisms, and even better (-1.79% Spatial and -0.9% Temporal Distribution) than Spatial Cloaking\(^9\) for black-box attacks.

\(^6\)With an exception for the Rounding mechanism because information lost by rounding data points cannot be reconstructed by an adversary.

\(^7\)Including data from rows 8-11 columns 1-4.

\(^8\)Including data from rows 8-11 columns 5-8.

\(^9\)Best of all analyzed LPPMs against black-box attacks.
posed a mechanism based on maximum-entropy as alternative Location-Based Services. Theodorakopoulos [5], instead, proposed an alternative approach that makes use of agility maneuvers to spatial-cloaking and geo-indistinguishability. We proposed an alternative approach that makes use of agility maneuvers for the manipulation of real location data.

VI. Conclusion

This paper has considered a new science of environment reconfiguration called system agility, by proposing agile maneuvers that have been evaluated against heuristics adoptable by cybercriminals for inference attacks on location data. We found out that agile maneuvers are more robust against white-box attacks resulting in a probability of success of only 2.68% on average, compared to an average of 56.92% when using state-of-the-art Location-Privacy Preserving Mechanisms. Future work should investigate the impact of agile maneuvers on legitimate uses of location data, and the users themselves.

V. Related Work

Krumm [2] examined location data gathered from volunteer subjects to quantify how well four heuristics can be used by an adversary to identify the subjects’ home locations. We extended Krumm’s work by analyzing how well three agility mechanisms work against both black-box and white-box attacks. Golle et al. [3] showed that obfuscation techniques are less effective if the subject’s home and work locations are known or deducible from external sources (i.e., online search engine). We extended their work by studying how well six previously known and three new protection mechanisms prevent black-box and white-box attacks aiming to identify victims’ home and work locations from location data.

Andrés et al. [4] proposed geo-indistinguishability, a mechanism to add controlled noise to the user’s location in order to obtain an approximate version of it to be sent to Location-Based Services. Theodorakopoulos [5], instead, proposed a mechanism based on maximum-entropy as alternative to spatial-cloaking and geo-indistinguishability. We proposed an alternative approach that makes use of agility maneuvers for the manipulation of real location data.

REFERENCES