POSTER: Compromising Cloaking-based Location Privacy Preserving Mechanisms with Location Injection Attacks

Lei Jin
School of Information Sciences
University of Pittsburgh
Pittsburgh, PA, USA
lej17@pitt.edu

Balaji Palanisamy
School of Information Sciences
University of Pittsburgh
Pittsburgh, PA, USA
bpalan@pitt.edu

James Joshi
School of Information Sciences
University of Pittsburgh
Pittsburgh, PA, USA
jjoshi@apitt.edu

ABSTRACT

Cloaking-based location privacy preserving mechanisms have been widely adopted to protect users' location privacy while traveling on road networks. However, a fundamental limitation of such mechanisms is that users in the system are inherently trusted and assumed to always report their true locations. Such vulnerability can lead to a new class of attacks called location injection attacks which can successfully break users' anonymity among a set of users through the injection of fake user accounts and incorrect location updates. In this paper, we characterize location injection attacks, demonstrate their effectiveness through experiments on real-world geographic maps and discuss possible defense mechanisms to protect against location injection attacks.

Categories and Subject Descriptors

H.2.7 [**Database Management**]: Database Administration—*Security, integrity, and protection*; H.2.8 [**Database Management**]: Database Applications—*Spatial databases and GIS*

General Terms

Experimentation, Security

Keywords

Location Cloaking, Location Privacy, Location k-Anonymity, Location Injection Attack

1. INTRODUCTION

Location privacy threats refer to the risks that an adversary can obtain unauthorized access to raw location data by locating a transmitting device and identifying the subject (person) using the mobile device. Examples of such risks include spamming users with unwanted advertisements, drawing sensitive inferences from victims' visits to clinics and doctors' offices and learning one's religious activities and political beliefs. Location privacy is a system-level capability of location-based systems, which control the access to location information at different spatial granularities and different temporal and continuity scales, rather than stopping all access to location information.

CCS'14, November 3–7, 2014, Scottsdale, Arizona, USA.
ACM 978-1-4503-2957-6/14/11.
http://dx.doi.org/10.1145/2660267.2662386.

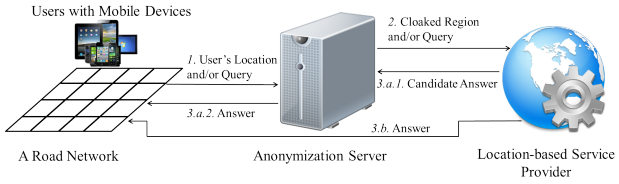

Figure 1: Trusted anonymizer architecture.

In the past, cloaking-based location privacy preserving mechanisms (CLPMs) have been proposed as one of the most effective location privacy preserving mechanisms for users traveling on the road networks [1, 2, 4]. As shown in Figure 1, when a user requests a location-based service (e.g. searching for the nearest coffee shop) from a Location-based Service Provider (LSP), he first sends the request to a trusted Anonymization Server (AZ) which launches a location cloaking algorithm to reduce the precision of the user's location and generates a cloaked region under the required granularity level. The AZ then sends the cloaked location to the LSP to obtain the required location-based service. Here, the LSP can be a potential adversary and can be either curious or malicious.

In general, any cloaking-based location privacy preserving mechanism (CLPM) guarantees the in-distinguishability of a given user among a set of other users. *Location k-Anonymity* [1] refers to the property that ensures that the location of a given subject (user) is indistinguishable from that of $k-1$ other users. In addition to *location k-Anonymity*, several extensions to the basic CLPMs have been proposed to strengthen the privacy guarantees including *POI* (*points of interest*) *l-Diversity* [1] which ensures the in-distinguishability of a user's location from a set of POIs and *Segment s-Diversity* [4] which guarantees the in-distinguishability of a user's location from a set of road segments. However, in all existing CLPMs, a fundamental limitation is that all users are inherently trusted by the AZ and assumed to always report their true locations. In this work, we show that such vulnerability can lead to a new class of attacks called location injection attacks which can successfully violate users' privacy in terms of in-distinguishability among a set of users. In this paper, we first characterize the location injection attack and then demonstrate its effectiveness for CLPMs through experiments on real-world geographic maps (Section 2). Finally, we discuss the potential solutions which can be utilized to identify and mitigate location injection attacks (Section 3).

2. LOCATION INJECTION ATTACK

In this section, we illustrate the location injection attack with an intuitive example and formally define the attack and propose the attack schemes.

Illustrative example: We consider the example shown in Figure 2 where we find six real users u_1, u_2, u_3, u_4, u_5 and u_6 traveling in a road network and an adversary creates six fake users fu_1, fu_2, fu_3, fu_4, fu_5 and fu_6 and reports their locations in the road segments around the road junction, $Jun1$. Assuming user u_1 has a *location k-Anonymity* requirement, $k^{u_1} = 6$ and without the presence of fake users, AZ would generate a cloaked region containing users u_1, u_2, u_3, u_4, u_5 and u_6. In this case, the probability of identifying the exact location of u_1 from that of others is 1/6. However, when the adversary launches a location injection attack, AZ may generate a cloaked region including segments $Seg1$ and $Seg3$ where there are only two real users (u_1, u_2) and four non-real attack users (fu_2, fu_4, fu_5, fu_6). Since the adversary (who is launching the attack) can distinguish fake users in the generated cloaked region, the probability of identifying u_1 from others is now reduced to 1/2, which violates the *location k-Anonymity* requirement of u_1. Hence, the adversary now has a higher probability to identify u_1's exact location (for e.g., u_1 could be traveling in the $Seg1$, $Seg2$, $Seg3$, $Seg4$ or $Seg11$ without the attack but when the location injection attack is launched, u_1 should be in either $Seg1$ or $Seg3$).

2.1 Attack Definition

We assume that there is a road network $G(V_G, E_G)$ where V_G represents the set of road junctions and E_G represents the set of road segments between the junctions. We consider that there is an authentic user *u* who travels in *G* while requesting the location service from a location-based service provider (LSP) through an Anonymization Server (AZ). *u* has a privacy setting k^u for *location k-Anonymity* and AZ guarantees k^u in the generated cloaked region. We also assume that the adversary is part of LSP who tries to violate *u*'s privacy setting of k^u. Here, we also assume that the adversary (LSP) knows *u*'s coarse location (l_u^0) before launching the attack. This information can be obtained by the cloaked regions that *u* used for his recent queries to LSP.

Adversary's Actions: The adversary intelligently manipulates a number of fake user locations to be similar and close to l_u^0 and as a result, AZ generates a cloaked region R^u for *u* and sends with associated users (including *u*) in R^u to the LSP. We define the set $U(R^u)$ as a set containing all users in R^u and a set $U(R^u)'$ denoting the set of fake users in R^u.

Successful Location Injection Attack: We say that user u is a victim of a *location injection attack*, when $|U(R^u)| - |U(R^u)'| < k^u$. Here $|U|$ indicates the number of users in a user set *U*.

We note that not every location injection attack is successful and we consider that the attack is successful only when the number of authentic users is less than that required. For example, in Figure 2, if the anonymity requirement k^{u_1} of the user u_1 is 2, then u_1's privacy requirement is not violated even under the location injection attack. We also note that a location injection can be targeted at multiple users simultaneously and the attack can be used to infer a targeted user's trajectory when the continuous location injection attacks for the targeted user are successful.

2.2 Attack Costs

In general, there are two types of costs associated with the adversary to launch location injection attacks:

- *Cost of creating non-real user account*: there is a cost to create non-real users and we define such cost as $cost_F(n)$ where *n* denotes number of non-real users created.

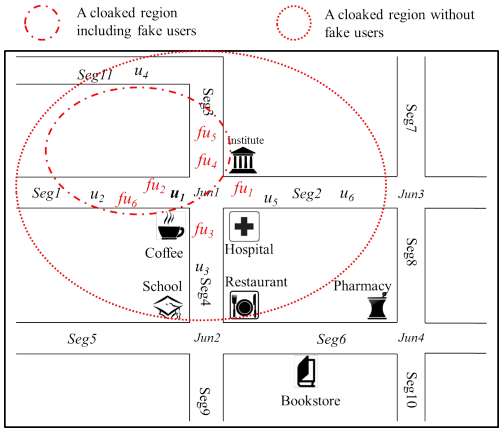

Figure 2: An instance of a location injection attack.

- *Cost of the anonymity service from AZ*: we assume that there is a cost for a user *u* whenever *u* requests the anonymity service from AZ. We assume the cost of the anonymity service for each user is equal and we denote it as $cost_u(m)$ where *m* indicates times the anonymity service is requested by user *u*.

Hence, the total cost $cost_A$ of a location injection attack can be calculated as the sum of the fake user creation cost and cost of users' anonymization requests:

$$\cos t_A = \cos t_F(n) + \sum_{i=1}^{i=n} \cos t_{u_i}(m_{u_i})$$

2.3 Attack Schemes

We define two types of attack schemes for location injection attacks namely:

- *Random Injection* (RI) *attack*: Given a targeted user *u*, an adversary can randomly choose a number of fake users and has no limitation on how their locations are manipulated. For instance, a fake user in the random injection attack may travel 10 miles within 10 seconds through location manipulation by the adversary and we assume that the AZ may not be aware of abnormal traveling activities.

- *Trajectory-based Injection* (TI) *attack*: In trajectory-based injection attacks, an adversary creates normal trajectories for fake users by simulating these trajectories similar to a targeted user's trajectory.

Intuitively, the attack cost of the RI attack is less than that of the TI attack since the fake user accounts can be significantly reused in the RI attack. Also, the RI attack is even more cost-effective when the adversary targets multiple users at the same time. Next, we experimentally evaluate and demonstrate the effectiveness of location injection attacks under these two attack schemes.

2.4 Attack Simulations

We use the GT Mobile simulator [3] to generate a trajectory of 5000 users moving in Northwest Atlanta regions of Georgia. We assume that all of these 5000 users are authentic. We implement the Anonymization Server (AZ) using the road network-aware *XStar* cloaking algorithm [4] to preserve users' privacy requirements in terms of *location k-Anonymity* and *segment s-diversity*. For each authentic user u, we set k^u as a randomly chosen value from 2 to

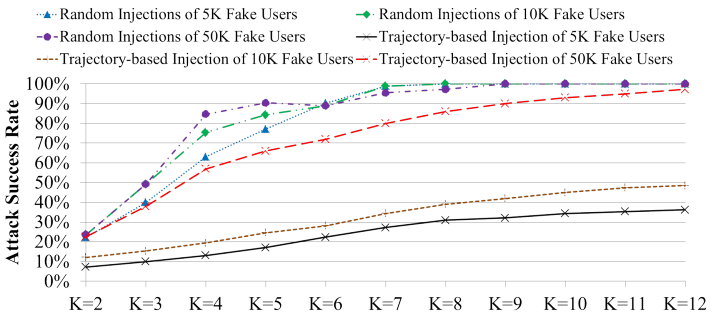

Figure 3: Attack results of both attack schemes.

12 and s^u (*Segment s-Diversity* [4]) is randomly chosen from 2 to 10. We assume that each user in the road is active and he or she requests the location-based service from the location-based service provider (LSP) every second. Out of the 5000 authentic users, we randomly choose 800 users in the road network as targets. For each target, the adversary conducts both the RI attacks and TI attacks every time a target has a location-based query. We evaluate the attack effectiveness by injecting 5000, 10000 and 50000 fake users for both attack schemes. We note that in our experiments, the total attack cost is the same for both the RI attack and TI attack in terms of the same number of fake users injected.

We use the attack success rate as the metric to show the effectiveness of the attacks. When a location injection attack is successful for a target, it indicates that a target's privacy requirement of *location k-Anonymity* is violated (less than $k-1$ authentic users in the cloaked region). Figure 3 shows the observed success rate of the attacks where X-axis represents the *location k-Anonymity* requirements of targeted users and Y-axis refers to the average attack success rate for the targeted users. For example, K=5 refers to a group of the targets who specify their privacy requirements of *location k-Anonymity* to be 5. As shown in the Figure, the average attack rate is around 90% in RI attacks when 5000 fake users are utilized. From the Figure 3, we can see that an adversary can have a more successful attack when a target has a larger value for *location k-Anonymity*. The average attack success rates also gradually increase with the increase in the number of injected fake users in both attack schemes. Additionally, we find that the average attack success rates of the RI attacks are usually much higher than those of the TI attacks in terms of the same number of fake users injected.

We also note that location injection attacks do not violate a user *Segment s-Diversity* s^u as the *XStar* cloaking algorithm typically ensures *segment s-diversity* independent of *location k-Anonymity*. However, a larger value of could possibly decrease the effectiveness of the attack. When there are more segments in a cloaked region, it is more likely to include many authentic users and thus, a cloaking algorithm without the *Segment s-Diversity* guarantee would be even more vulnerable to location injection attacks.

3. DISCUSSIONS

In this section, we discuss some possible defense mechanisms and our on-going research on developing solution techniques to identify and mitigate location injection attacks. Intuitively, if an adversary has limited privilege to arbitrarily assign locations to fake users, an adversary needs more efforts and a higher cost to launch location injection attacks. Therefore, a straight-forward defense approach would be to identify and blacklist users who do not follow a normal trajectory. In this case, it is more difficult for an adversary to reutilize a part of fake users whose previous locations are not close to the current location of a targeted user. This approach can significantly mitigate *Random Injection* attacks. However, it is still ineffective for the *Trajectory-based Injection* attacks as fake users in this attack follow a normal trajectory.

To defend against *Trajectory-based Injection* attacks, a potential solution is to detect possible suspicious traveling activities of the fake users (e.g., circle a location for a long time) and blacklist them. When a user has a suspicious trajectory, the user can be blacklisted and may not be included in the effective anonymity set of the cloaked region. However, with such an approach, the key challenge is to identify and effectively characterize all suspicious traveling activities on a road network which may be difficult in practice. Also, the detection approach needs to be updated whenever a new characteristic of a fake user activity is identified. In addition, we also note that some authentic users may get blacklisted as part of false negatives (e.g., an authentic user circles a stadium to find a parking slot and the detection mechanism may incorrectly identify him as a fake user.) and will experience some form of denial of service.

In our ongoing and future work, we are working on developing a trust-based defense mechanism for identifying and blacklisting suspicious user activities on road networks. The objective of the solution is to minimize the attack effectiveness by significantly increasing the cost of an effective attack while incurring zero or minimal impact on the service quality to the authentic users.

4. REFERENCES

[1] B. Bamba, L. Liu, P. Pesti, and T. Wang. Supporting anonymous location queries in mobile environments with privacygrid. In *Proceedings of the 17th International Conference on World Wide Web*, WWW '08, pages 237–246, New York, NY, USA, 2008. ACM.

[2] M. F. Mokbel, C.-Y. Chow, and W. G. Aref. The new casper: Query processing for location services without compromising privacy. In *Proceedings of the 32Nd International Conference on Very Large Data Bases*, VLDB '06, pages 763–774. VLDB Endowment, 2006.

[3] P. Pesti, B. Bamba, M. Doo, L. Liu, B. Palanisamy, and M. Weber. Gtmobisim: A mobile trace generator for road networks, 2009.

[4] T. Wang and L. Liu. Privacy-aware mobile services over road networks. *Proc. VLDB Endow.*, 2(1):1042–1053, Aug. 2009.