On the Anonymization of Differentially Private Location Obfuscation

Yusuke Kawamoto
National Institute of Advanced
Industrial Science and Technology (AIST)
Tsukuba, Japan

Takao Murakami
National Institute of Advanced
Industrial Science and Technology (AIST)
Tokyo, Japan

***Abstract*—Obfuscation techniques in location-based services (LBSs) have been shown useful to hide the concrete locations of service users, whereas they do not necessarily provide the anonymity. We quantify the anonymity of the location data obfuscated by the planar Laplacian mechanism and that by the optimal geo-indistinguishable mechanism of Bordenabe et al. We empirically show that the latter provides stronger anonymity than the former in the sense that more users in the database satisfy k-anonymity. To formalize and analyze such approximate anonymity we introduce the notion of asymptotic anonymity. Then we show that the location data obfuscated by the optimal geo-indistinguishable mechanism can be anonymized by removing a smaller number of users from the database. Furthermore, we demonstrate that the optimal geo-indistinguishable mechanism has better utility both for users and for data analysts.**

I. Introduction

Location-based services (LBSs) have been increasingly employed in a variety of applications, including navigation, resource-trucking, recommendation, advertising, games, and authentication. One of the popular applications has been to discover interesting locations from collected location data and provide them for third parties. When the providers of LBSs publish some geographic locations of users, the accurate locations may reveal private information, such as home addresses, health conditions, and political orientation.

To prevent or mitigate the privacy breach, many location obfuscation techniques have been proposed to hide accurate locations of users while providing their approximate information used in LBSs. For example, the *dummy location insertion* [1] generates $k-1$ dummy points and makes a user's location indistinguishable among a set of k locations, which provides k-*anonymity*. The *spacial cloaking technique* [2] chooses a sufficiently large region that includes k indistinguishable locations to achieve k-anonymity. The *location perturbation technique* [3] adds to each location a controlled random noise and guarantees *differential privacy*, independently of any side information that an adversary may possess.

Such perturbation techniques have been developed to construct more practical mechanisms for location obfuscation. The *planar Laplacian mechanism* [4] satisfies *geo-indistinguishability*, an extended notion of differential privacy to the Euclid distance. The *optimal geo-indistinguishable mechanism* [5] minimizes the quality loss caused by the perturbation while preserving geo-indistinguishability.

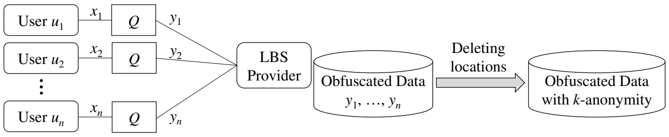


Fig. 1. Overview of the proposed method. Each user u_i obfuscates a location x_i using a mechanism Q and sends the obfuscated location y_i to the LBS provider, which anonymizes the collected data to publish them.

Although these geo-indistinguishable mechanisms hide the concrete locations, no prior work has investigated the relationships between geo-indistinguishability and anonymity to our knowledge. In this paper, we show geo-indistinguishability does not guarantee to provide k-anonymity. This means that the location data obfuscated by geo-indistinguishable mechanisms might be vulnerable to re-identification attacks (e.g., [6], [7]) for instance when the LBS provider shares the obfuscated data with a malicious data analyst. Moreover, such leakage of user identity information can be efficiently detected and quantified using an automated tool such as [8], [9].

In this work we empirically explore the relationships among obfuscation, anonymity, and utility for users and for data analysts in geo-indistinguishable location obfuscation. In particular, we propose a method for effectively anonymizing the obfuscated data by deleting some data before publishing them to third parties. The overview of the method is shown in Fig. 1.

The contributions of this paper are summarized as follows:

- We evaluate the anonymity of the location data obfuscated by two location obfuscation mechanisms: PL (the planar Laplacian mechanism) and OptQL (the optimal geo-indistinguishable mechanism). We empirically show that OptQL satisfies stronger anonymity than PL.
- We propose the notion of (κ, α)-*asymptotic anonymity*, which generalizes k-anonymity to an approximate anonymity of sampled users.
- We show that the location deletion method, which simply removes the locations of the users who do not satisfy

This work was partially supported by JSPS KAKENHI Grant Number JP17K12667, JP16K16069, and by JSPS and Inria under the project LOGIS of the Japan-France AYAME Program.

k-anonymity, makes the location dataset k-anonymous while preserving ε-geo-indistinguishability. In particular, we demonstrate that OptQL requires to delete a smaller number of users than PL to achieve k-anonymity.

- We demonstrate by experiments that the utility for users and for data analysts is better in OptQL than in PL.

II. Preliminaries

For a finite set $\mathcal{S}$, we denote by $\#\mathcal{S}$ the number of elements in S, and by $\mathbb{D}\mathcal{S}$ the set of all probability distributions over $\mathcal{S}$.

A. Obfuscation Mechanism

In this work we consider a number n of users each reporting some rough information y on his single geographic location x to an LBS (location-based service) provider while keeping the exact location x hidden from the provider. To compute an obfuscated location y, each user uses a *location obfuscation mechanism* that adds a certain noise to x and outputs it as y.

Formally, let $\mathcal{X}$ be a finite set of all possible locations of the users, and $\mathcal{Y}$ be a finite set of all (possibly fake) locations reported by the users. Then a *location obfuscation mechanism* (or simply an *obfuscater*) is a probabilistic algorithm $Q : \mathcal{X} \rightarrow \mathbb{D}\mathcal{Y}$ that, given an original location x, outputs a *reported location* y. We denote by Q_{xy} the conditional probability that the mechanism Q outputs y given input x.

The probability distribution of the original locations is represented by the prior π over $\mathcal{X}$, and the prior probability of a location x is denoted by π_x.

B. Geo-indistinguishability

Geo-indistinguishability [4] is a notion of location privacy that can be regarded as a variant of local differential privacy [10] in which the privacy budget ε is multiplied by the Euclidean distance $d(x, x')$ between locations x and x'.

Definition 1 (ε-geo-indistinguishability): Given $\varepsilon \geq 0$, an obfuscation mechanism Q provides *ε-geo-indistinguishability* if for any inputs $x, x' \in \mathcal{X}$ and any output $y \in \mathcal{Y}$, we have:

$$Q_{xy} \leq e^{\varepsilon d(x,x')} Q_{x'y}.$$

Then the difference between Q_{xy} and $Q_{x'y}$ are proportional to the distance between x and x'. This implies that geo-indistinguishability allows an adversary to infer approximate information about the original location (e.g., a user is in Paris), but hides the exact location (e.g., home address) from her. By relaxing the privacy requirements in this way, the amount of noise added to the location can be significantly reduced (compared to local differential privacy [10]). Consequently, geo-indistinguishability is useful to implement practical LBSs such as the POI (point of interest) retrieval [4].

C. Planar Laplacian (PL) Mechanism

The *planar Laplacian (*PL*) mechanism* [4] is an example of the mechanism providing geo-indistinguishability. It generates a random noise according to a two-dimensional Laplace distribution, and obfuscates an original location x by adding the noise to x. In this paper we use a variant of the planar Laplacian mechanism, which outputs a symbol "$\perp$" when the obfuscated location is outside the area of interest $\mathcal{X}$.

Formally, the variant planar Laplacian mechanism $Q^{\mathsf{PL}} : \mathcal{X} \rightarrow \mathbb{D}(\mathcal{X} \cup \{\perp\})$ is defined by:

$$Q^{\mathsf{PL}}_{xy} = \begin{cases} \frac{1}{c} \cdot e^{-\varepsilon d(x,y)} & \text{(if } y \in \mathcal{X}\text{)} \\ 1 - \frac{1}{c} \cdot \sum_{y' \in \mathcal{X}} e^{-\varepsilon d(x,y')} & \text{(if } y = \perp\text{)}, \end{cases}$$

where $c = \max_x \sum_{y' \in \mathcal{X}} e^{-\varepsilon d(x,y')}$. Intuitively, c is selected to have the best utility by preventing unnecessarily frequent outputs of $\perp$.

Proposition 1: Q^{PL} satisfies ε-geo-indistinguishability.

Proof: Q^{PL} can be seen as a cascade of the standard planar Laplacian (that does not output $\perp$) and the post-processing algorithm that maps each $y \notin \mathcal{X}$ to $\perp$. It is easy to see that by the triangle inequality, the standard planar Laplacian satisfies ε-geo-indistinguishability. Since differential privacy is immune to post-processing, Q^{PL} provides ε-geo-indistinguishability. □

D. Optimal Geo-indistinguishability (OptQL) Mechanism

The planar Laplacian mechanism is efficiently computable while the utility of the reported location may not be optimal. For this reason, Bordenabe *et al.* [5] propose an *optimal geo-indistinguishable location obfuscation mechanism* OptQL that given a privacy budget ε, minimizes the quality loss (QL) that is defined as the expected value of the Euclidean distance, i.e.,

$$QL(\pi, Q, d) = \textstyle\sum_{x,y} \pi_x Q_{xy} d(x, y).$$

The mechanism OptQL can be obtained by solving a linear optimization problem that minimizes $QL(\pi, Q, d)$ while satisfying ε-geo-indistinguishability. However, the computational complexity of this optimization is in $O(\#\mathcal{X}^3)$. To reduce this to $O(\#\mathcal{X}^2)$, they show an approximation technique based on a spanning graph of the set of locations. See [5] for details.

E. k-Anonymity

The notion of *k-anonymity* [11] of a user ensures that the user cannot be distinguished from at least $k - 1$ other users being at the same location. More formally, for a positive integer k, we say that the users at a location y are k-anonymous if $n(y) \geq k$ where $n(y)$ is the number of the users who report y as their locations. We also say that a dataset of locations satisfies k-anonymity if for every location y in the dataset, the users at y are *k-anonymous*. In this definition k-anonymity depends only on the users that have the lowest level of anonymity, and does not take the other users into account.

III. Anonymization of Obfuscated Location Data

In this section we address some limitations in the definition of k-anonymity and introduce two anonymity notions that generalize k-anonymity. The first notion measures an obfuscater's capability of anonymization independently of the number n of sampled users. The second notion extends the first one to take into account the fact that different users in the dataset may have different levels of anonymity. Finally, we present a simple solution for enhancing the anonymity of the obfuscated data while preserving geo-indistinguishability.

A. Limitations in the Definition of k-Anonymity

k-anonymity is not always useful to evaluate the level of anonymity in the presence of sampled users.

First, k-anonymity in the context of location privacy depends on the number n of the LBS's users in a sample data, and does not solely express an obfuscater Q's capability of anonymization. For instance, if the number n of sampled users increases then k-anonymity tends to hold for a larger value of k (roughly proportionally to n) for the same π and Q. In other words, k-anonymity is not defined as a property of (π, Q) independently of the number n of sampled users.

Second, different users in the dataset may have different anonymity levels, whereas k-anonymity of the dataset depends only on the users that have the lowest level of anonymity. Hence k-anonymity is not expressive enough to take into account the different anonymity levels of the other users.

B. κ-Asymptotic Anonymity

To overcome the first limitation described in Section III-A, we introduce a notion that expresses an obfuscater Q's capability of anonymization independently of the number n of sampled users. Intuitively, for a $\kappa \in [0,1]$, we define the notion of *κ-asymptotic anonymity* as an extension of k-anonymity where for any sufficiently large number n of users, each user is indistinguishable from roughly $n \cdot \kappa - 1$ other users.

Formally, this notion is defined using the probability $p(y)$ that the obfuscation mechanism Q outputs y as follows.

Definition 2 (κ-asymptotic anonymity): Given a threshold $\kappa \in [0,1]$, the users at a location y are *κ-asymptotically anonymous* if $p(y) > \kappa$. Given a prior $\pi \in \mathbb{D}\mathcal{X}$ and an obfuscater $Q : \mathcal{X} \to \mathbb{D}\mathcal{Y}$, we say that (π, Q) provides *κ-asymptotic anonymity* if for all $y \in \mathcal{Y}$, $p(y) > 0$ implies $p(y) > \kappa$, where $p(y) = \sum_x \pi_x Q_{xy}$.

Note that κ itself can be computed from π and Q independently of n. When (π, Q) provides κ-asymptotic anonymity, the number of users required to achieve k-anonymity is roughly given by $\frac{k}{\kappa}$.

Example 1 (Anonymity of the prior and posterior): Let us formalize the asymptotic anonymity before/after applying a mechanism Q. The prior π provides $(\min_x \pi_x)$-asymptotic anonymity while (π, Q) provides $(\min_y \sum_x \pi_x Q_{xy})$-asymptotic anonymity[1]. To achieve k-anonymity before (resp. after) applying Q, the number of users should be roughly $\frac{k}{\min_x \pi_x}$ (resp. $\frac{k}{\min_y \sum_x \pi_x Q_{xy}}$).

For a large number n of users, we can compute an approximate maximum value of κ from the sample by $\min_y \hat{p}(y) = \min_y \frac{n(y)}{n}$, which converges to κ quickly as shown in Fig. 5.

As we will see in Section IV-B1, κ-asymptotic anonymity (resp. k-anonymity) holds only for small values of κ (resp. k). This implies that the obfuscation mechanism does not necessarily provide anonymity to all users although it hides the exact original locations in terms of geo-indistinguishability.

C. (κ, α)-Asymptotic Anonymity

Similarly to k-anonymity, the definition of κ-asymptotic anonymity also suffers from the second limitation described in Section III-A. To evaluate the different levels of anonymity of different users, we introduce another notion that relaxes κ-anonymity by allowing some rate α of errors. Roughly speaking, the new notion expresses that given a sample data with n users, at least $n(1-\alpha)$ users are $n\kappa$-anonymous.

Definition 3 ((κ, α)-asymptotic anonymity): Let $p(y) \stackrel{\text{def}}{=} \sum_x \pi_x Q_{xy}$. Given a $\kappa \in [0,1]$ and an acceptable error rate $\alpha \in [0,1]$, (π, Q) provides (κ, α)*-asymptotic anonymity* if

$$\frac{\sum_{y:p(y)>\kappa} p(y)}{\sum_{y:p(y)>0} p(y)} \geq 1 - \alpha.$$

This notion can be used to roughly estimate the utility loss in anonymizing the location data. When there are n users in the dataset, at most $n\alpha$ users are not $n\kappa$-anonymous. If we remove the locations data of these users, then the dataset will satisfy $n\kappa$-anonymity while the utility of the dataset deteriorates proportionally to the number $n\alpha$ of deleted users.

D. Location Deletion Method (Del) for k-Anonymity

As explained so far, ε-geo-indistinguishable mechanisms are useful to hide the exact locations from the LBS provider, whereas they may not be able to provide k-anonymity of the obfuscated location data. When the LBS provider wishes to publish such obfuscated data to third parties, a simple solution to achieve k-anonymity is what we call the *location deletion method* Del, i.e., to delete the obfuscated locations that do not satisfy k-anonymity. Then the modified database satisfies k-anonymity while preserving ε-geo-indistinguishability thanks to the immunity to the post-processing.

More specifically, given a threshold κ, the minimum number of users that should be removed is approximately given by:

$$n\alpha_{\min} = n \cdot \left(\frac{\sum_{y:0<p(y)<\kappa} p(y)}{\sum_{y:p(y)>0} p(y)} \right),$$

where $p(y) = \sum_x \pi_x Q_{xy}$. When Q is a Laplacian mechanism, then all locations occur with non-zero probabilities, and thus the approximate number of deleted users is $n \cdot \sum_{y:p(y)<\kappa} p(y)$. We will demonstrate the effect of this combination of obfuscation and anonymization by experiments in Section IV.

IV. Experimental Evaluation

In this section we empirically compare the two obfuscation mechanisms PL and OptQL, and illustrate how the location deletion method Del enhances the anonymity of obfuscated data and affects the utility for users and for data analysts.

[1] Remarkably, the asymptotic anonymity contrasts with the Bayes-vulnerability (aka. converse of the Bayes risk [12]) in quantitative information flow. Instead of minimization, the prior/posterior Bayes-vulnerabilities are respectively $\max_x \pi_x$ and $\sum_y \max_x \pi_x Q_{xy}$, and represent the probabilities of an adversary's correctly guessing x in one attempt before/after observing y.

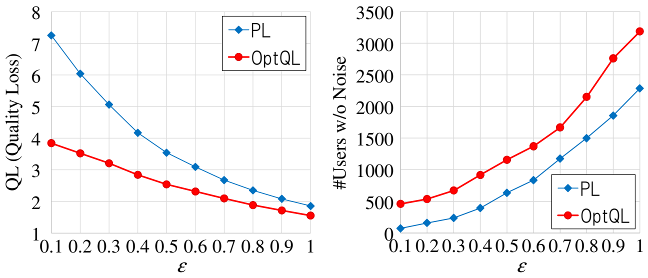


Fig. 2. Trade-offs between the privacy budget ε and the utility for users. As a utility the graph on the left uses QL (quality loss), and the graph on the right uses the number of users that remain at the same regions after obfuscation. As for PL, we excluded the users who report $\perp$ as their location.

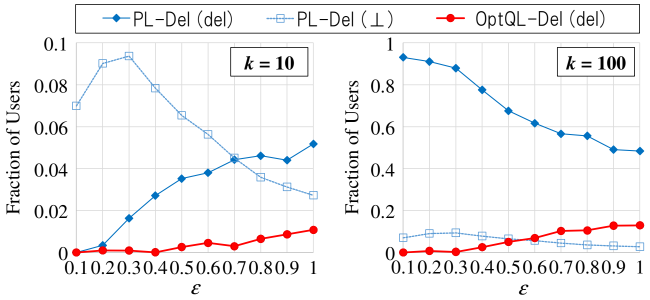


Fig. 3. Trade-offs between the privacy budget ε and the utility for data analysts. The y-axis represents the fraction of deleted users necessary to satisfy κ-anonymity (PL−Del (del), OptQL−Del (del)), and that of users who output $\perp$ as reported locations (PL−Del ($\perp$)), where $\kappa = 6.689 \times 10^{-4}$ ($k = 10$) on the left and $\kappa = 6.689 \times 10^{-3}$ ($k = 100$) on the right.

A. Experimental Set-up

We performed experiments using the Foursquare dataset (Global-scale Check-in Dataset) [13]. This dataset includes 33278683 location check-ins by 266909 users all over the world. In our experiments, we used the data in Manhattan, which consists of location check-ins by 14951 users. We assumed that each user u_i obfuscated a single location x_i using an ε-geo-indistinguishable obfuscation mechanism Q, and sent the obfuscated location y_i to the LBS provider.

We divided Manhattan into 20×20 regions with regular intervals. Let $\mathcal{X}$ be the set of these regions, and π be the empirical distribution of the 14951 users' locations over $\mathcal{X}$. We defined the distance $d(x, x')$ between two regions x and x' by the Euclidean distance between their central points. Here we normalized the distance so that the distance between two adjacent regions is one.

As an obfuscation mechanism Q, we employed the planar Laplacian mechanism PL (in Section II-C) and the Optimal geo-indistinguishable mechanism OptQL (in Section II-D). In OptQL, we solved the optimization problem[2] that minimizes QL while satisfying ε-geo-indistinguishability using the linear programming solver linprog in MATLAB. For both PL and OptQL, we set the privacy budget ε to be 0.1 to 1, which have been widely used in the literature [14].

After obtaining all obfuscated regions $y_1, y_2, \ldots, y_n$, we applied the location deletion method Del to remove the regions that do not satisfy k-anonymity (where k is 10 or 100). We denote by PL−Del (resp. OptQL−Del) the application of PL (resp. OptQL) post-processed by Del.

B. Experimental Results

We show the experimental results on anonymity and utility.

1) k-anonymity before anonymization: By experiments we found that unless we add much noise, the obfuscation does not provide k-anonymity, i.e., $k = 1$ for a user. Specifically, $k = 1$ is provided by PL for $\varepsilon \geq 0.4$ and by OptQL for $\varepsilon \geq 0.2$.

2) Utility for users: In Fig. 2 we compare OptQL with PL in terms of the utility for users. Specifically, we evaluated the quality loss, i.e., the average Euclidean distance $d(x_i, y_i)$ between the original region x_i and the obfuscated region y_i. We also evaluated the number of users who remain at the same region after obfuscation, i.e., $x_i = y_i$. As shown in Fig. 2, for a larger ε, smaller noise is added, hence both PL and OptQL have better utility for users; They decrease the quality loss, and increase the number of users remaining at the same regions. The results also demonstrate that OptQL outperforms PL in terms of the utility for users. This is because OptQL chooses locations that minimize the expected distance, which also makes more users remain at the same regions.

3) Utility for data analysts: In Fig. 3 we compare OptQL with PL in terms of the utility for data analysts. The graphs show the ratio of deleted users for $\kappa = 6.689 \times 10^{-4}$ ($k = 10$) on the left and for $\kappa = 6.689 \times 10^{-3}$ ($k = 100$) on the right. As for PL we also show the ratio of users reporting $\perp$ as obfuscated regions (indicated as PL−Del ($\perp$)).

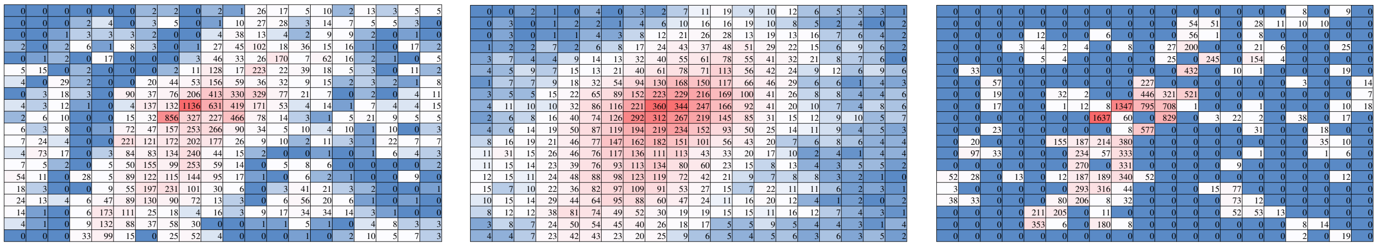
According to these graphs, the ratio of deleted users is significantly smaller in OptQL−Del than in PL−Del. To see this in detail, we present the maps of Manhattan that plot the density of the user locations without noise (Fig. 4a), and of those obfuscated by PL (Fig. 4b) and by OptQL (Fig. 4c).

In Fig 4b we see that the planar Laplacian PL spreads the population over the whole map. This is because PL uniformly draws an angle (from $[0, 2\pi)$) to which it maps each location. For $\varepsilon \approx 0$, the reported regions are distributed almost uniformly. Hence for a small value of k, only a few obfuscated regions need to be deleted to achieve k-anonymity (Fig. 3 on the left), whereas for a large value of k, most of the obfuscated locations need to be deleted (Fig. 3 on the right).

In contrast to PL, *the optimal geo-indistinguishable mechanism* OptQL *concentrates more users in the crowded regions* as shown in Fig. 4c. To see this in detail, we note that for a more crowded region x, the prior probability π_x is larger. Since OptQL tries to minimize $\sum_{x,y} \pi_x Q_{xy} d(x, y)$, if π_x is larger then OptQL chooses a region y with a smaller $d(x, y)$, i.e., closer to x. Hence the users located in the crowded regions tend not to move by the obfuscation. Conversely, the users outside the crowded regions tend to move to one of the closest crowded regions that provide geo-indistinguishability.

Owing to this concentration, OptQL provides (κ, α)-

[2] In OptQL we set the dilation factor to be $\delta = 1.09$.



(a) When the users reported their original locations.

(b) When the users reported the locations obfuscated by PL ($\varepsilon = 1$).

(c) When the users reported the locations obfuscated by OptQL ($\varepsilon = 1$).

Fig. 4. The maps of Manhattan that plot the numbers of users having reported the regions as their (original/obfuscated) locations.

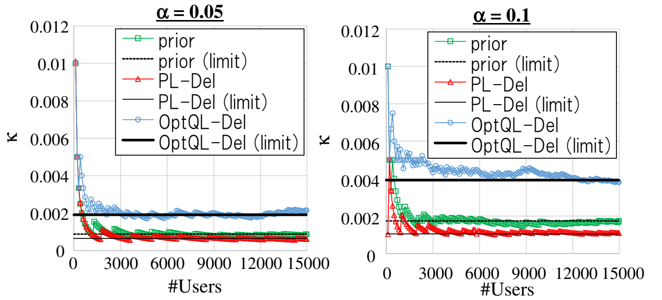


Fig. 5. Relationship between the number of users and the level κ of the asymptotic anonymity for $\alpha = 0.05$ on the left and for $\alpha = 0.1$ on the right. As for PL−Del we excluded $\perp$ from the computation of κ for clarity.

TABLE I
THE LEVEL κ OF ASYMPTOTIC ANONYMITY CONVERGES TO THE FOLLOWING VALUES WHEN INCREASING THE NUMBER OF USERS.

	$\alpha = 0.05$	$\alpha = 0.1$
prior (without noise)	8.7×10^{-4}	1.7×10^{-3}
after applying PL−Del	6.4×10^{-4}	1.0×10^{-3}
after applying OptQL−Del	1.9×10^{-3}	3.9×10^{-3}

asymptotic anonymity with a smaller error rate α. For instance, in OptQL, only 161 users do not satisfy 10-anonymity ($\alpha = 0.011$), whereas in PL, 773 users do not ($\alpha = 0.052$). This means that OptQL−Del removes a smaller number of users than PL−Del, and thus has a better utility for data analysts.

To sum up OptQL−Del is more effective than PL−Del in terms of the utility both for users and for data analysts while providing ε-geo-indistinguishability and k-anonymity.

4) Convergence of the empirical value of κ: In Fig. 5 we show how the empirically computed value of κ converges to the value displayed in Table I when increasing the number n' of users. In the experiments we uniformly sampled a subset (of size n') from the original dataset, applied each mechanism with $\varepsilon = 1$, and computed the maximum κ such that $n'(1-\alpha)$ users satisfy $n'\kappa$-anonymity (for $\alpha = 0.05, 0.1$). These graphs imply that κ is (roughly) independent of n' and thus κ-asymptotic anonymity can be seen as a property of the prior and obfuscater. Therefore κ is useful to learn that given a different number n of sampled users, the dataset roughly satisfies $n\kappa$-anonymity.

V. Conclusion

We have empirically evaluated the anonymity of the location data obfuscated by PL and by OptQL, and shown that OptQL provides stronger anonymity than PL in the sense that it requires to remove a fewer users to achieve k-anonymity. To analyze this formally, we have introduced the notion of (κ, α)-asymptotic anonymity. We have also demonstrated that OptQL has better utility for users and for data analysts.

In future work we plan to develop a utility-optimal obfuscater satisfying geo-indistinguishability and anonymity. We will also explore rigorous foundations of obfuscation based on statistics, and relationships with quantitative information flow.

References

[1] H. Kido, Y. Yanagisawa, and T. Satoh, “Protection of location privacy using dummies for location-based services,” in *Proc. of ICDE Workshops*, 2005, p. 1248.

[2] M. Gruteser and D. Grunwald, “Anonymous usage of location-based services through spatial and temporal cloaking,” in *Proc. of MobiSys*. USENIX, 2003.

[3] A. Machanavajjhala, D. Kifer, J. M. Abowd, J. Gehrke, and L. Vilhuber, “Privacy: Theory meets practice on the map,” in *Proc. of ICDE*. IEEE, 2008, pp. 277–286.

[4] M. E. Andrés, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, “Geo-indistinguishability: differential privacy for location-based systems,” in *Proc. of CCS’13*. ACM, 2013, pp. 901–914.

[5] N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, “Optimal geo-indistinguishable mechanisms for location privacy,” in *Proc. of CCS’14*, 2014, pp. 251–262.

[6] Z. Montazeri, A. Houmansadr, and H. Pishro-Nik, “Achieving perfect location privacy in markov models using anonymization,” in *Proc. of ISITA’16*, 2016, pp. 355–359.

[7] T. Murakami, “A succinct model for re-identification of mobility traces based on small training data,” in *Proc. of ISITA’18*, 2018, to appear.

[8] T. Chothia, Y. Kawamoto, and C. Novakovic, “A tool for estimating information leakage,” in *Proc. of CAV’13*, 2013, pp. 690–695.

[9] ——, “LeakWatch: Estimating information leakage from java programs,” in *Proc. of ESORICS’14 Part II*, 2014, pp. 219–236.

[10] J. C. Duchi, M. I. Jordan, and M. J. Wainwright, “Local privacy and statistical minimax rates,” in *Proc. of FOCS’13*, 2013, pp. 429–438.

[11] L. Sweeney, “k-anonymity: A model for protecting privacy,” *Int. Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 5, pp. 557–570, 2002.

[12] K. Chatzikokolakis, C. Palamidessi, and P. Panangaden, “On the Bayes risk in information-hiding protocols,” *J. of Comp. Security*, vol. 16, no. 5, pp. 531–571, 2008.

[13] D. Yang, D. Zhang, and B. Qu, “Participatory cultural mapping based on collective behavior data in location based social networks,” *ACM Transactions on Intelligent Systems and Technology*, vol. 7, no. 3, pp. 30:1–30:23, 2015.

[14] J. Hsu, M. Gaboardi, A. Haeberlen, S. Khanna, A. Narayan, B. C. Pierce, and A. Roth, “Differential privacy: An economic method for choosing epsilon,” in *Proc. of CSF’14*, 2014, pp. 398–410.