IEEE Open Journal of Vehicular Technology

Received 14 January 2024; accepted 1 February 2024. Date of publication 9 February 2024;
date of current version 28 March 2024. The review of this article was coordinated by Editor Ping Wang.

Digital Object Identifier 10.1109/OJVT.2024.3364184

A Framework for Tradeoff Between Location Privacy Preservation and Quality of Experience in Location Based Services

TIANYI FENG [1], ZHIXIANG ZHANG [1] (Graduate Student Member, IEEE), WAI-CHOONG WONG [1] (Life Senior Member, IEEE), SUMEI SUN [2] (Fellow, IEEE), AND BIPLAB SIKDAR [1] (Senior Member, IEEE)

[1]Department of Electrical, Computer Engineering, National University of Singapore, Singapore 119077
[2]Institute for Infocomm Research, Agency for Science, Technology, Research, Singapore 138632

CORRESPONDING AUTHOR: TIANYI FENG (e-mail: fengtianyi@u.nus.edu)

This work was supported by the National Research Foundation, Singapore and Infocomm Media Development Authority under its Future Communications Research and Development Programme under Grant FCP-NUS-TG-2022-001.

ABSTRACT Location-based services find a number of applications in vehicular environments such as navigation, parking, infortainment etc. However, the disclosure of vehicles' location information raises multiple privacy issues. To balance the tradeoff between privacy and utility, this paper proposes a framework to preserve users' location privacy while delivering the desired quality of experience (QoE). The proposed framework allows users to quantify the data utility while accessing location-based services under different privacy levels through the QoE metric. The privacy analysis of the proposed framework is provided under two adversary models. Finally, the effectiveness of the proposed framework is demonstrate using the real-world "Dianping" review dataset.

INDEX TERMS Differential privacy, location-based services, location privacy, quality of experience.

I. INTRODUCTION

Location-based services (LBS) are gaining increasing popularity in vehicular environments. Typical LBS in vehicular scenarios include mapping and navigation, parking services, weather forecast, nearby Points of Interest (POI), infotainment, and location-based social networking. LBS bring convenience to drivers and passengers, but also cause serious privacy concerns. If an untrusted application service provider is able to access a user's location information continuously, more sensitive or private information can be extracted, such as the user's home address, occupation, relationships and even health conditions [1].

With the privacy concerns associated with LBS [2], users may be unwilling to disclose their true whereabouts to other entities. To address these privacy issues, various location privacy preservation mechanisms (LPPM) have been proposed. Commonly used approaches for privacy preservation include anonymization [3], obfuscation [4], position dummies [5], encryption [6], and mix-zones [7]. In addition, differential privacy [8] is a mathematical construct to provide provable privacy to any individual whose data is in a statistical database.

The implementation of the various approaches for privacy preservation mentioned above results in a loss of utility then the data is used by various applications and location based services. As a result, while implementing mechanisms for privacy, it is also important to ensure data utility or application service quality. Therefore, a tradeoff between location privacy and data utility should be provided by any location privacy preserving framework. However, in previous works, no rigorous utility functions have been considered while evaluating such a tradeoff. Some researchers have introduced a general or empirical utility function to express the data utility cost [9], [10]. To comprehensively address this problem, we propose two performance metrics, Quality of Service (QoS) and Quality of Experience (QoE), to quantify the data utility with either objective or subjective measures. QoS is a common metric and has been used for quantification of service performance

in areas such as communications and software engineering. In contrast, QoE measures the overall experience or the satisfaction by end-users, which is both subjective and objective [11]. With such metrics, the gap between location privacy and application service quality can be filled. We also compare the performance of QoE and QoS to determine which is the more representative metric to quantify the data utility in our defined scenario. Additionally, we evaluate the performance of our framework under two adversarial scenarios, benign and malicious, each with different objectives and capabilities.

In this paper, we propose a location privacy preserving framework based on differential privacy, and propose a framework for assessing the tradeoff between privacy and data utility. The main contributions of this paper can be summarized as follows.

1) We present a LBS system solution to preserve location privacy with differential privacy and characterize the tradeoff between location privacy and data utility.
2) We introduce two metrics, QoS and QoE, to quantify the data utility. To the best of our knowledge, this is the first analytical framework that establishes a QoE model mathematically and logically for a LBS system and analyzes the tradeoff between location privacy and QoE.
3) In addition to vehicular environments, the proposed LPPM can be implemented in any localization system to preserve users' location privacy.
4) The performance of the proposed LPPM is evaluated under two adversarial models. The first one is a honest-but-curious adversary such as a LBS provider with unlimited access to a user's data while the second one only has access to the user's interaction with the LBS server over the network and uses a Hidden Markov Model (HMM) to infer the users' real locations as the hidden states with given observable pseudo-locations.

The rest of the paper is organized as follows. The related works on LPPMs and QoE are introduced in Section II. Section III presents a new definition for differential privacy aimed at location based services that is used in our framework. The system model, performance metrics, and the design of the proposed LPPM are described in Section IV. We establish a QoE model based on a real-life dataset and quantify the tradeoff between location privacy and QoE in Section V-A. Two adversary models with different capabilities are proposed in Section VI. The results related to evaluation of proposed framework are presented in Section VII. Section VIII concludes this paper.

II. RELATED WORK

Location privacy preservation has been a very active research topic. Shokri et al. [12] formalized the problem as a Bayesian Stackelberg game and proposed a game-theoretic framework for protecting users' location privacy. They also proposed metrics to quantify location privacy and considered the adversary's prior knowledge to obtain the optimal user-centric LPPM, which can anticipate the inference attack and concurrently satisfy the service quality requirement [13]. Andrés et al. proposed the notion of 'Geo-Indistinguishability' to guarantee no leakage of the user's exact location by releasing approximate location information [14]. Their mechanism can protect the user's location privacy within a radius corresponding to a privacy level. In particular, geo-indistinguishability is a modified and generalized version of differential privacy. Differential privacy is a well-known concept providing a constraint on mechanisms to preserve an individual's privacy [8]. Moreover, both Shokri et al. and Andrés et al. used linear programming techniques to achieve optimal privacy and minimize the loss of service quality. Bordenabe et al. proposed an approach that uses a spanning graph to approximate distances between locations to reduce the total number of constraints in the linear program from cubic to quadratic since the linear optimization is time-consuming and computationally demanding [15].

There are also many other location privacy preserving methods, such as k-anonymity [16], obfuscation, encryption, position dummies, mix-zones, and their combination. The Privacy-Preserving Paradigm-driven framework for indoor Localization (P^3-Loc) [17] employed k-anonymity and differential privacy approaches to guarantee both the user's and the location server's privacy. P^3-Loc took the advantage of the fact that the localization process of most IPSs consist of two phases: the online phase for estimating location and the offline phase for measuring information. P^3-Loc perturbed and cloaked the transmitted data in both phases. Position dummies [18] protect a user's true location by generating and sending diversified fake positions to the location server together with the real location. Shankar et al. [19] proposed the 'SybilQuery' scheme to generate Sybil queries based on decentralized and autonomous k-anonymity to ensure dummies cannot be discriminated from the user's true location. 'PShare' is a cryptography-based approach proposed by Wernke et al. [20] to solve the problem in non-trusted systems. They utilized the concept of multi-secret sharing by splitting up the user's location information into shares and distributing them to multiple non-trusted location servers. Beresford et al. [21] proposed a mix-zone-based approach to protect identities in defined areas named mix zones. Within a mix zone, all the users must hide their identities and cannot send any location updates. Furthermore, all the users' identities need to be mixed by exchanging or changing pseudonyms so that an adversary cannot track users continuously.

Most of the above-mentioned approaches neglected data utility quantification metrics or adversary's possible counteractivities while designing LPPMs. These two parts cannot be disregarded since data utility is a crucial and necessary metric in a location privacy preserving framework and the robustness of LPPMs needs to be evaluated with appropriate adversary models [22]. To address this open problem, in this paper, we propose two metrics to quantify data utility and a solution to find the tradeoff between location privacy and application

service quality. In addition, we design two adversary models and validate the effectiveness and efficiency of our proposed framework with under these models.

QoE is a measurement of the perception or satisfaction of a customer's experience with a service [23], which can be applied to diverse service areas [24], [25]. Compared with QoS, QoE is more subjective and takes more factors of influence into consideration. Moreover, current studies combine different factors to quantify the QoE, such as linear regression [26], exponential [27], decision tree [28], etc. In [29], several existing QoE functions are employed together to evaluate a mobile video streaming algorithm's performance. QoE has also been applied in some location-related scenarios. For example, [30] proposed and created a quantitative link between localization accuracy and QoE in museums and exhibitions. In this paper, we first introduce QoE into a location privacy preserving framework and then use QoE to evaluate the application service performance of the proposed LPPM by conducting a case study with real-word data.

III. LBS CENTRIC DIFFERENTIAL PRIVACY

In this section, we present a new definition for differential privacy that takes the data utility into consideration. Next, the justification for the choice of the mechanism for achieving differential privacy is presented.

A. DEFINITIONS

The core concept of differential privacy is that for two adjacent datasets with only one different record, the probability of obtaining the same results by querying these two databases should be quite close [33].

Definition 1 (Differential Privacy): A randomized algorithm $\mathcal{M}$ gives (ϵ, δ)-differential privacy if for all $\mathcal{S} \subseteq Range(\mathcal{M})$ and for all datasets D_1, D_2 such that $\|D_1 - D_2\|_1 \leq 1$,

$$\Pr\{\mathcal{M}(D_1) \in S\} \leq \exp(\epsilon) \times \Pr\{\mathcal{M}(D_2) \in S\} + \delta \quad (1)$$

where $\|D_1 - D_2\|_1$ is the distance between D_1 and D_2 to measure how many records differ in the two datasets. If $\delta = 0$, the randomized mechanism $\mathcal{M}$ provides ϵ-differential privacy, which is the strictest definition.

The differential privacy level ϵ represents how much perturbation is needed for a specific privacy level. More specifically, if a user wants to achieve a higher degree of privacy protection, the differential privacy level ϵ should be set smaller or close to 0. In addition, it is the ℓ_1 sensitivity [34] since only one important parameter determines the degree of privacy level.

Following the principle above, we can derive a generalized notion of differential privacy for localization.

Definition 2 (ϵ-Differential Privacy for Localization): A location privacy preserving mechanism $\mathcal{M}$ satisfies ϵ-differential privacy, if for any input true location l_i and any output pseudo location l_i', the following holds

$$\frac{\Pr\{\mathcal{M}(l_i) = l_i'\}}{\Pr\{\mathcal{M}(l_j) = l_i'\}} \leq e^{\epsilon}, \forall j \neq i, \quad (2)$$

where l_j is any location on the map other than the input true location l_i.

Definition 2 implies that a constraint related to differential privacy level ϵ should restrict the probabilities of the same pseudo location given different real locations. Then, the user's real locations cannot be distinguished from the received pseudo locations.

While preserving location privacy in a LBS system, some practical considerations need to be taken into account, and it is unreasonable to obfuscate the user's real position to anywhere on the map, e.g., an unrealistic remote position. Thus, we need location service quality thresholds for ensuring the data utility of the generated pseudo locations. With the location service quality threshold, we modify the traditional definition of differential privacy with a constraint.

Definition 3 (Loc-correlated privacy): A randomized algorithm $\mathcal{M}$ gives ϵ loc-correlated privacy if for any two input true locations l_i, l_j and any output pseudo location l_i', all $j \neq i$ and $l_i' \in Range(\mathcal{M}(l_i)) \cap Range(\mathcal{M}(l_j))$,

$$\Pr\{\mathcal{M}(l_i) = l_i'\} \leq \exp(\epsilon) \times \Pr\{\mathcal{M}(l_j) = l_i'\} \quad (3)$$

where the distributions of generated pseudo locations of l_i and l_j have overlapping areas.

These three definitions are the theoretical foundations of the proposed framework. Definition 1 represents the original definition of differential privacy. Definition 2 is the generalized notion of differential privacy in the localization scenario. Definition 3 involves a practical constraint to be applied in a LBS system. In particular, our framework follows Definition 3 to protect the user's location privacy.

B. MECHANISMS FOR ACHIEVING DIFFERENTIAL PRIVACY

Typical methods to realize differential privacy are adding a random noise, such as Laplacian noise and Gaussian noise [31], and randomized response [32].

Definition 4 (Laplace Mechanism): Given any function $f: D \to \mathbb{R}$, the Laplace mechanism $\mathcal{M}_L$ is defined as

$$\mathcal{M}_L(D) = f(D) + Lap\left(\frac{\Delta f}{\epsilon}\right). \quad (4)$$

The Laplace mechanism achieves ϵ-differential privacy. Alternatively, the Gaussian mechanism can achieve (ϵ, δ)-differential privacy, which scales to ℓ_2-sensitivity [34].

Definition 5 (Gaussian Mechanism): For any $\delta \in (0, 1)$, given any function $f: D \to \mathbb{R}$, the Gaussian mechanism $\mathcal{M}_G$ is defined as

$$\mathcal{M}_G(D) = f(D) + (Y_1, \ldots, Y_k) \quad (5)$$

where Y_i are i.i.d. random variables drawn from the Gaussian distribution $N(0, \sigma^2)$ and $\sigma \geq \frac{\Delta_2 f \sqrt{2 \ln(\frac{1.25}{\delta})}}{\epsilon}$.

C. MECHANISM SELECTION BY ENTROPY

The principle of selecting the better mechanism for location privacy preservation is to determine which mechanism is more challenging for the adversary to infer a user's real position. Therefore, we use the concept of entropy [35] to quantify the uncertainty.

Definition 6 (Entropy): Given a random variable X, Shannon entropy $H(X)$ is defined as

$$H(X) = \mathbb{E}[I(X)] = \mathbb{E}[-\log(\Pr(X))] \tag{6}$$

where $\mathbb{E}$ is the expected value operator, and I is the information content of X.

Then, we can compute the entropy of Laplace mechanism H_L as

$$H_L = \ln\left(\frac{2\Delta f}{\epsilon}\right) + 1. \tag{7}$$

The entropy of Gaussian mechanism H_G can be computed as

$$H_G = \ln\left(\sigma\sqrt{2\pi}\right) + \frac{1}{2} \geq \ln\left(\frac{2\Delta_2 f}{\epsilon}\sqrt{\pi \ln\left(\frac{1.25}{\delta}\right)}\right) + \frac{1}{2}. \tag{8}$$

The entropy difference ΔH between the Laplace and Gaussian mechanisms is defined as

$$\begin{aligned}\Delta H &= H_G - H_L \\ &= \ln\left(\frac{2\Delta_2 f}{\epsilon}\sqrt{\pi \ln\left(\frac{1.25}{\delta}\right)}\right) + \frac{1}{2} - \ln\left(\frac{2\Delta f}{\epsilon}\right) - 1 \\ &= \ln\left(\frac{\Delta_2 f}{\Delta f}\sqrt{\pi \ln\left(\frac{1.25}{\delta}\right)}\right) - \frac{1}{2}\end{aligned} \tag{9}$$

where $\frac{\Delta_2 f}{\Delta f} \leq 1$. If $\Delta H < 0$, we should select the Laplace mechanism. To make $\Delta H < 0$, δ should satisfy the constraint $\delta > 0.526$. In addition, the differential privacy with ℓ_2-sensitivity can be represented as

$$\Pr\{\mathcal{M}(D_1) \in S\} \leq \exp(\epsilon) \times \Pr\{\mathcal{M}(D_2) \in S\} + \delta, \tag{10}$$

which implies that the Gaussian mechanism can achieve (ϵ, δ)-differential privacy or ϵ-differential privacy with a parameter at least $1 - \delta$. It also shows that the privacy guarantee of Gaussian noise is weaker than Laplacian noise and the larger the δ, the weaker the privacy guarantee of the Gaussian mechanism. Thus, we can conclude that δ should always be small and cannot be larger than 0.526 to ensure better location preserving performance of Gaussian mechanism. With this choice of δ, (9) will always be less than zero, namely

$$\Delta H = H_G - H_L < 0. \tag{11}$$

Based on this inequality, we can conclude that the Gaussian mechanism provides less uncertainty than the Laplace mechanism, which means that the pseudo-locations from different real locations can be distinguished more accurately with the Gaussian mechanism. Therefore, for location privacy preservation, implementing the Laplace mechanism is the more appropriate choice.

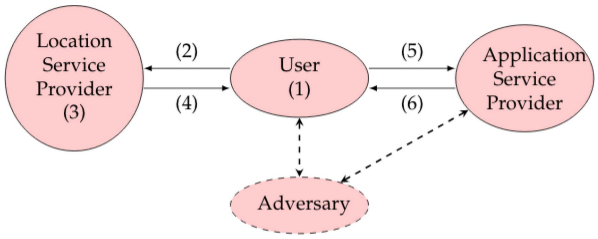


FIGURE 1. System model.

IV. PROPOSED FRAMEWORK

A. SYSTEM MODEL

We propose a location privacy preserving framework for scenarios where an adversary may compromise a LBS system to try to access the user's location information. Taking into consideration the probable presence of an adversary, there are four entities in the proposed system model as shown as Fig. 1. In step 1, a user first selects the privacy level ϵ in terms of the specific privacy requirement. Then, the user requests for the localization service from the location service provider in step 2 and forwards the received location information to the application service provider to access LBS in step 5. In our system model, the location service provider estimates the user's location and generates a pseudo location with the built-in LPPM in step 3. Finally, the application service provider returns the service results to the user in step 6. With the above procedure, this system can provide localization service and LBS for users without disclosing their true locations.

B. ADVERSARY MODEL

Different adversaries may have various levels of capability and access to information to launch successful attacks. We broadly classify the adversaries into two groups: internal and external. An external adversary can only monitor the interaction of the user and the service provider and tries to infer the user's true location from these exchanges. An internal adversary (e.g., the LBS application provider) has access to all queries, results and the user's selection from the results, and is thus more powerful. We defer the full description of the adversary model to Section VI in order to relate the attack strategies of the attackers to the specifics of the proposed LPPM.

C. METRICS

1) LOCATION PRIVACY

In order to protect users' location privacy, their real location information should be kept private by using pseudo-locations instead, and a user-predefined privacy level ϵ determines the obfuscation degree of true locations.

Definition 7 (Location Privacy ψ): For user i, let the true location l_i be denoted as (x_i, y_i) and the pseudo-location l_i' be

denoted as (x_i', y_i'). The location privacy ψ_i can be derived as the Euclidean distance between l_i and l_i' as $d(l_i, l_i')$:

$$\psi_i = d(l_i, l_i') = \sqrt{(x_i - x_i')^2 + (y_i - y_i')^2}. \tag{12}$$

This definition of location privacy is only related to the distance between the true location and the pseudo-location. In practice, to ensure localization service quality, we need to introduce two thresholds $r_{\min}$ and $r_{\max}$ to constrain the scope of pseudo locations and these two radii $r_{\min}$ and $r_{\max}$ are relevant to the location privacy ψ as $r_{\min} \leq \psi_i = d(l_i, l_i') \leq r_{\max}$. Within this range, we can compute an average location privacy over the new transformed Laplacian distribution.

Definition 8 (Average Location Privacy $\bar{\psi}$): For user i, the mean of Laplacian noise is r_{Lap} and the Laplacian distribution can be derived as

$$f(r) = \epsilon \cdot \exp(-2\epsilon |r - r_{\text{Lap}}|). \tag{13}$$

With the constraint on r, the probability distribution of pseudo locations changes to a new distribution as

$$g(r) = \frac{f(r)}{\int_{r_{\min}}^{r_{\max}} f(r)dr}, r \in (r_{\min}, r_{\max}). \tag{14}$$

The average location privacy $\bar{\psi}_i$ over the new distribution is

$$\bar{\psi}_i = \int_{r_{\min}}^{r_{\max}} r g(r) dr = \int_{r_{\min}}^{r_{\max}} r \frac{f(r)}{\int_{r_{\min}}^{r_{\max}} f(r)dr} dr. \tag{15}$$

2) APPLICATION SERVICE QUALITY

Application service quality is an essential metric to evaluate the performance of pseudo-locations in a location privacy preserving LBS system. The more privacy to protect, the farther the pseudo-location is, and the user acquires lower application service quality. We introduce QoS and QoE as two metrics to quantify the application service quality in this paper.

Definition 9 (QoS): For user i with true location l_i:(x_i, y_i) and pseudo-location l_i':(x_i', y_i'), the application service provider offers LBS within a radius r_{LBS} from the center l_i' instead of the center l_i since the user queries the LBS with his or her pseudo-location. We define the overlapping region area of these two circles as S_{op} and the area of application service circle as S_{LBS}. Then, the QoS after the obfuscation is defined as

$$QoS = \frac{S_{\text{op}}}{S_{\text{LBS}}}. \tag{16}$$

The detailed computation method is discussed in Section V-A.

Definition 10 (QoE): QoE is a multi-dimensional metric with objective factors $\boldsymbol{O}$ and subjective factors $\boldsymbol{S}$ of specific services for each user, while using a LBS system. A general QoE function is defined as

$$QoE = F(\boldsymbol{O}, \boldsymbol{S}), \tag{17}$$

where $\boldsymbol{O}$ represents objective factors (e.g., screen resolution in a cinema) and $\boldsymbol{S}$ is a set of subjective factors (e.g., flavor score for a restaurant).

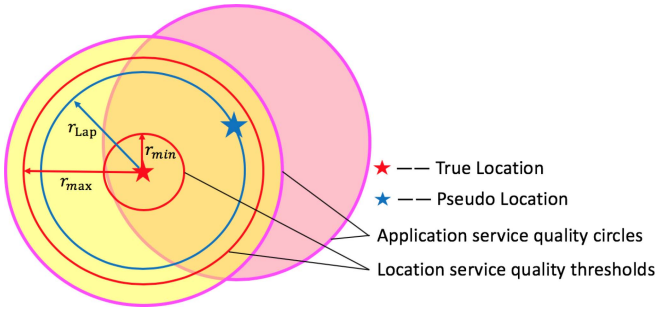


FIGURE 2. Problem scenario.

D. PROPOSED LPPM

In our scenario, the user requests for pseudo-locations from the location service provider to access LBS and pseudo-locations should be generated with specific privacy level. This scenario can be modeled as a problem as shown in Fig. 2. The red and blue stars represent the true location and one possible pseudo location, respectively. The red circle shows the range of the user's acceptable localization results, which we call the location service quality threshold. Similarly, the purple circle illustrates the user's expected range of the service results, which we call the application service quality requirement. In addition, the blue circle's radius shows the Laplacian noise added to the true location and the orange overlapping region is the remaining useful application service area, which is defined as the application service quality.

In our proposed framework, we follow the principle of Loc-correlated privacy to preserve users' location privacy. The detailed procedure of our LPPM model is introduced as follows. For any given true location (x_i, y_i) and any privacy level ϵ, we can compute the Laplacian noise r_{Lap} added to the true location [14] within the location service quality threshold $(r_{\min}, r_{\max})$ by

$$r_{\text{Lap}} = \left| -\frac{1}{\epsilon} \left(W_{-1} \left(\frac{p-1}{e} \right) + 1 \right) \right|, r_{\min} \leq r_{\text{Lap}} \leq r_{\max} \tag{18}$$

where $p = rand(1)$ is uniformly distributed in the interval [0,1] and W_{-1} is the LambertW function. Since $p \in [0, 1]$, r_{Lap} has its own range $(0, \frac{1}{\epsilon})$ and $(r_{\min}, r_{\max})$ should be within this range. The relationship among these parameters can be summarized as $0 \leq r_{\min} \leq r_{\text{Lap}} \leq r_{\max} \leq \frac{1}{\epsilon}$.

The next step is to choose the direction by setting a random angle $\theta = rand(1) \cdot 2\pi$. Finally, we can determine the pseudo location as

$$\begin{cases} x_i' = x_i + r_{\text{Lap}} \cdot \cos(\theta) \\ y_i' = y_i + r_{\text{Lap}} \cdot \sin(\theta). \end{cases} \tag{19}$$

E. PROOF OF LOC-CORRELATED PRIVACY

In this section, we prove that the proposed LPPM satisfies ϵ Loc-correlated privacy. The conclusion we need to prove is

$$\frac{\Pr\{\mathcal{M}(l_i) = l_i'\}}{\Pr\{\mathcal{M}(l_j) = l_i'\}} \leq e^{\epsilon}, \tag{20}$$

where $j \neq i$, and i, j should ensure that l_i' is within the location service quality thresholds of l_i and l_j. This can be transformed as

$$\frac{\Pr\{\mathcal{M}(l_i) = l_i'\}}{\Pr\{\mathcal{M}(l_j) = l_i'\}} \leq \frac{\max(\Pr\{\mathcal{M}(l_i) = l_i'\})}{\min(\Pr\{\mathcal{M}(l_j) = l_i'\})}. \tag{21}$$

Since $0 \leq r_{\min} \leq r_{\text{Lap}} \leq r_{\max} \leq \frac{1}{\epsilon}$, $\Pr\{\mathcal{M}(l_i) = l_i'\}$ will be maximized when $r_{\text{Lap}} = \frac{r_{\max}+r_{\min}}{2}$, and $\Pr\{\mathcal{M}(l_i) = l_i'\}$ will be minimized when $r_{\text{Lap}} = r_{\min}$. Thus,

$$\begin{cases} \max(\Pr\{\mathcal{M}(l_i) = l_i'\}) = \int_0^{\frac{1}{\epsilon}} \epsilon \cdot \exp(-2\epsilon|r - \frac{1}{2\epsilon}|)dr \\ \qquad\qquad = -\frac{1}{2}(e^{-1} - 1), \\ \min(\Pr\{\mathcal{M}(l_j) = l_i'\}) = \int_0^{\frac{1}{\epsilon}} \epsilon \cdot \exp(-2\epsilon r)dr \\ \qquad\qquad = -\frac{1}{2}(e^{-2} - 1). \end{cases} \tag{22}$$

Also, $\frac{-\frac{1}{2}(e^{-1}-1)}{-\frac{1}{2}(e^{-2}-1)} = 1 - \frac{1}{1+e} < 1 \leq e^{\epsilon}$. Then, we have

$$\frac{\Pr\{\mathcal{M}(l_i) = l_i'\}}{\Pr\{\mathcal{M}(l_j) = l_i'\}} \leq \frac{\max(\Pr\{\mathcal{M}(l_i) = l_i'\})}{\min(\Pr\{\mathcal{M}(l_j) = l_i'\})} < 1 \leq e^{\epsilon}. \tag{23}$$

Thus, our proposed mechanism satisfies the principle of Loc-correlated privacy.

V. TRADEOFF BETWEEN DATA PRIVACY AND UTILITY

In this section, we evaluate the privacy-utility tradeoff using the metrics that quantify data utility while protecting location privacy with the LPPM presented in Section IV-D.

A. QOS

With the pseudo-location (x_i', y_i'), two application service quality threshold circles can be drawn to illustrate the service quality performance. The area of the service quality threshold circle can be computed as $S_{\text{LBS}} = \pi r_{\text{LBS}}^2$, and the overlapping region area can be computed as

$$\begin{aligned} S_{\text{op}} &= 2(S_{\text{sector}} - S_{\text{triangle}}) \\ &= 2\left[\pi r_{\text{LBS}}^2 \left(\frac{2\arccos\left(\frac{r_{\text{Lap}}}{2r_{\text{LBS}}}\right)}{2\pi}\right) - \left(\frac{r_{\text{Lap}}}{2}\sqrt{r_{\text{LBS}}^2 - \left(\frac{r_{\text{Lap}}}{2}\right)^2}\right)\right]. \end{aligned} \tag{24}$$

Using (24), we can compute the ratio of the overlapping region area over the whole area of the application service quality threshold circle with only one pseudo-location as

$$\begin{aligned} QoS_N &= \frac{S_{\text{op}}}{S_{\text{LBS}}} \\ &= \frac{2\left[\pi r_{\text{LBS}}^2 \left(\frac{2\arccos\left(\frac{r_{\text{Lap}}}{2r_{\text{LBS}}}\right)}{2\pi}\right) - \left(\frac{r_{\text{Lap}}}{2}\sqrt{r_{\text{LBS}}^2 - \left(\frac{r_{\text{Lap}}}{2}\right)^2}\right)\right]}{\pi r_{\text{LBS}}^2} \\ &= \frac{2\arccos\left(\frac{r_{\text{Lap}}}{2r_{\text{LBS}}}\right)}{\pi} - \frac{r_{\text{Lap}}\sqrt{1 - \left(\frac{r_{\text{Lap}}}{2r_{\text{LBS}}}\right)^2}}{\pi r_{\text{LBS}}}. \end{aligned} \tag{25}$$

Using the probability distribution functions from (13) and (14), the average service quality over the whole transformed probability distribution can be derived as

$$\begin{aligned} Avg\_QoS &= \int_{r_{\min}}^{r_{\max}} QoS_N \cdot g(r)dr \\ &= \int_{r_{\min}}^{r_{\max}} \left(\frac{2\arccos\left(\frac{r}{2r_{\text{LBS}}}\right)}{\pi} - \frac{r\sqrt{1 - \left(\frac{r}{2r_{\text{LBS}}}\right)^2}}{\pi r_{\text{LBS}}}\right) \\ &\quad \cdot \frac{\epsilon \cdot \exp(-|r - r_{\text{Lap}}|2\epsilon}{\int_{r_{\min}}^{r_{\max}} \epsilon \cdot \exp(-|r - r_{\text{Lap}}|2\epsilon dr} dr. \end{aligned} \tag{26}$$

As we have shown in our previous paper [22], an increase in privacy level will degrade the application service quality.

B. QOE

1) DATASET

In contrast to QoS, QoE is a more subjective score based on the user's real personal experience and we need real users' experience data to train the QoE model. Therefore, we use the Dianping review dataset [36], which includes user reviews and detailed business information from a well-known Chinese review website. The Dianping website collects and records users' real ratings and reviews for specific businesses after visiting and using their services. The Dianping review dataset contains 3,605,300 reviews of 510,071 users towards 209,132 businesses. We pre-processed the dataset to extract the overall ratings, flavor scores, environment scores, service scores, and longitudes and latitudes of the restaurants in Beijing, China. With this dataset, we can define a QoE model in the scenario of accessing the LBS of requesting for nearby restaurant recommendations while preserving location privacy.

2) QOE MODEL

We define a general function, $QoE = F(\boldsymbol{O}, \boldsymbol{S})$, as the QoE function, where $\boldsymbol{O}$ is a set of objective factors and $\boldsymbol{S}$ is a set of subjective factors. Obviously, the set of objective and subjective factors depends on the specified scenario and datasets. Therefore, we train a specific QoE model with the real Dianping dataset. The following process of training a specific QoE model applies to other scenarios and datasets as well by adjusting the corresponding objective and subjective factors.

From the Dianping review dataset, we assume that the user's overall rating of the restaurant is the user's QoE value to mark the experience. Our QoE model is composed of several factors, such as flavor, environment, service and distance. In terms of the plots between the QoE rating and other factors shown in Fig. 3, we can see that QoE approximately follows a linear relationship with flavor, environment and service. Thus, we use multiple linear regression to train a QoE function with these three factors as

$$\begin{aligned} QoE &= \omega_0 + \omega_1 \times f + \omega_2 \times e + \omega_3 \times s \\ &= 1.413 + 0.656 \times f + 0.051 \times e + 0.207 \times s \end{aligned} \tag{27}$$

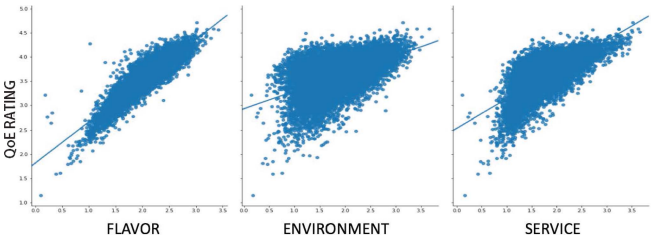


FIGURE 3. Relationship between rating and factors.

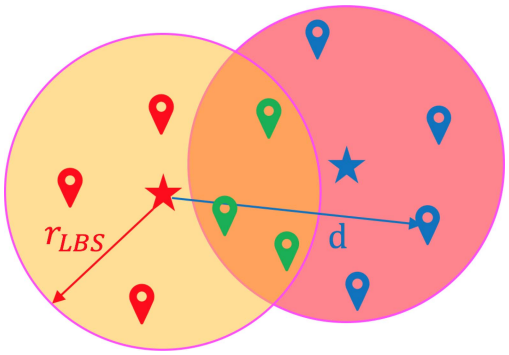


FIGURE 4. QoE model.

where QoE represents the overall rating, f is the flavor factor, e is the environment factor, and s is the service factor. The coefficient of each factor is obtained by using multiple linear regression for the Dianping review dataset. To evaluate this QoE function obtained through multiple linear regression, we compute the R-squared value, which is the coefficient of determination. The R-squared value is between 0 to 1 and the higher the value, the better the model fits the data [37]. In general, if $R^2 > 0.4$, the regression model has good imitative effect. In our QoE function, $R^2 = 0.8877$, so the model fits the data quite well.

Since we want to define a QoE model related to location privacy preservation, we incorporate the distance factor into the QoE function. As the scenario under consideration is the user requesting for nearby restaurant recommendations, the larger the distance between the user's location and the restaurant's location, the lower the QoE that the user obtains. To ensure that the user's original and perturbed application service threshold circles have an overlapping region, the maximal distance between the user and restaurant is $3r_{\text{LBS}}$. The QoE function with the distance factor is defined as

$$QoE_{\text{dis}} = (1.413 + 0.656f + 0.051e + 0.207s) \times \left(1 - \frac{d}{3r_{\text{LBS}}}\right) \tag{28}$$

where d represents the distance factor. The coefficients are obtained by using multiple linear regression for the Dianping review dataset presented in (27). Since users are requesting for nearby restaurant recommendations, the QoE should be inversely proportional to the distance. Also, the distance can never exceed $3r_{\text{LBS}}$. Thus, we propose the QoE function with distance factor as (27) multiplied by $(1 - \frac{d}{3r_{\text{LBS}}})$.

To compare QoE values based on true locations and pseudo-locations within the application service quality threshold, we normalize the QoE as

$$QoE_{\text{avg}} = \frac{1}{N}\sum_{1}^{N} QoE_{\text{dis}} \tag{29}$$

where QoE_{avg} is the average QoE within the LBS requirement.

Fig. 4 illustrates the proposed QoE model. The red and blue stars represent the true location and one possible pseudo location, respectively. The purple circles with the center of the red and blue stars show the application service quality requirement and the radii are the same as r_{LBS}. In addition, the red, green and blue spots illustrate the points of interest (POIs) of the user. If the user tries to access the LBS with his/her true location as the red star, the green and red POIs will be reported to him/her as the service results. However, the green and blue POIs will be reported to the user, if he/she accesses the LBS with the pseudo location, namely the blue star. Therefore, the green POIs are the overlapping service results and we can introduce a notion of QoE Loss to show the application service performance change.

Definition 11 (QoE Loss): For user i with real location l_i and pseudo location l_i', the user's QoE loss between accessing the LBS with the true location and the pseudo location can be defined as

$$QoE_{\text{Loss}} = QoE_{\text{avg\_real}} - QoE_{\text{avg\_pseudo}} \tag{30}$$

where $QoE_{\text{avg\_real}}$ is the average QoE of the POIs when the LBS is accessed with the real location, namely the average QoE value of the red and the green POIs to the red star, and $QoE_{\text{avg\_pseudo}}$ is the average QoE for the POIs when the LBS is accessed with the pseudo location, namely the average QoE value of the blue and the green POIs to the red star.

3) RELATIONSHIP BETWEEN QOE AND LOCATION PRIVACY

Before we try to find the tradeoff between QoE and location privacy, we need to characterize the relationship between QoE and location privacy. Since location privacy in the proposed framework is related to the differential privacy level, we obtain the mathematical relationship between QoE and differential privacy level ϵ instead. It is worth mentioning that the smaller the differential privacy level, the greater the location privacy that users can obtain. In our mechanism, we add Laplacian noise to achieve differential privacy and the probability density function (PDF) of the Laplace distribution is

$$p(x) = \frac{1}{2b}\exp\left(-\frac{|x|}{b}\right), b = \frac{\Delta f}{\epsilon}. \tag{31}$$

The expected estimation error $\mathbb{E}(err)$ of perturbation distances can be computed as

$$\mathbb{E}(err) = \sum p(x_i)\|x_i' - x_i\|_1$$

$$= \int_{-\infty}^{+\infty} |x| \frac{1}{2b} \exp\left(-\frac{|x|}{b}\right) dx = b = \frac{\Delta f}{\epsilon} \quad (32)$$

where x_i is the original value and x_i' is the perturbed value. Then, we can take the derivative with respect to the differential privacy level ϵ as

$$\frac{d\mathbb{E}(err)}{d\epsilon} = -\frac{\Delta f}{\epsilon^2} < 0. \quad (33)$$

Therefore, the expected estimation error $\mathbb{E}(err)$ decreases when the differential privacy level ϵ increases. If the application service quality threshold r_{LBS} stays the same, QoE will increase when $\mathbb{E}(err)$ decreases. Thus, QoE increases as the differential privacy level increases. In other words, users can enjoy the service with better experience quality, but their location privacy may be less protected. Above all, we proved the monotonicity between QoE and differential privacy level.

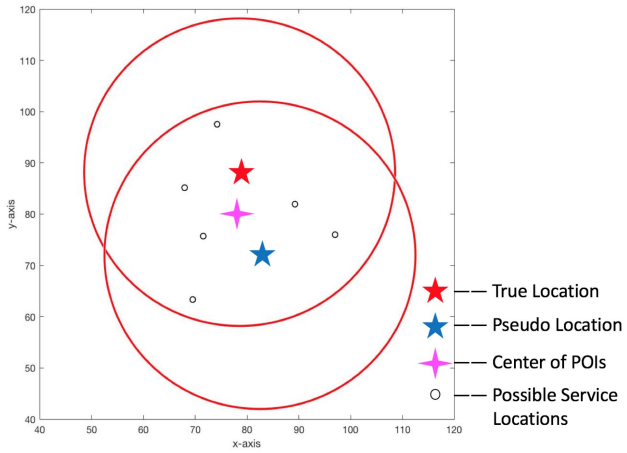


FIGURE 5. Possible points of interest for services.

VI. ADVERSARY MODELS

It is critical and necessary for a privacy preserving framework to evaluate its ability to protect the user's privacy against an adversary's inference. However, different adversaries with different intentions may have varying prior knowledge to design and execute their attacks. Therefore, we we consider two adversary models with different capabilities to evaluate the robustness of our proposed framework. For both scenarios, we assume that any entity who wants to obtain or infer the user's information without the user's permission can be considered as an adversary. The first case considers the LBS service provider as a honest-but-curious adversary who intends to infer the user's true location, e.g., to improve the application service quality. In the second case, the adversary is an external entity that tries to extract the user's location-related sensitive information for some harmful or commercial reasons.

A. INSIDER ADVERSARY

In this case, we assume that the adversary is the application service provider attempting to infer an user's true locations with the intent to improve service quality. Thus, the adversary can receive all the user's pseudo-locations directly and knows other information such as the user's POIs that the user browses among all the suggested service results (such as nearby restaurants).

The adversary can derive an user's true positions by dividing the whole region G into N squares and treating all the locations in the same grid as the grid center's coordinate. For instance, the adversary needs to transfer the received the pseudo location $l_i' : (x_i', y_i')$ to the corresponding grid $g_i : (x_{g_i}, y_{g_i})$. The probability of received pseudo locations in this grid $p(l_{g_i}')$ and the probability of possible points of interest for services in this grid $p(s_{g_i})$ can be calculated since the adversary has the history of all the pseudo locations. Then, the conditional probabilities of all the service selections given l_i' can be computed as

$$p(s_{g_j}|l_{g_i}') = \frac{p(s_{g_j}, l_{g_i}')}{p(l_{g_i}')}, \forall j \in N, \quad (34)$$

and $p(s_{g_j}, l_{g_i}')$ can be determined according to the recorded history.

It is reasonable to assume that users are only interested in the services satisfying their demands. This implies that users only browse the advised service results inside the overlapping region shown in Fig. 5. We also assume that the true location is distributed as a Laplacian distribution around the center of POIs l_c, as indicated by the purple cross in Fig. 5. The probability distribution of the possible true locations can be derived as

$$p(l_k) = 2\epsilon \cdot \exp(-|l_k - l_c| \cdot 2\epsilon). \quad (35)$$

In addition, there are some locations where users will never (or are extremely unlikely to) occur, such as in the middle of a river. Then, the probability distribution of the possibility and rationality of users' occurrence is

$$p(k) = \begin{cases} 0, & \text{if} \quad k \in F \\ 1, & \text{otherwise} \end{cases} \quad (36)$$

where F is the set of impossible locations. Finally, the probability of possible inferred true locations while receiving the pseudo location l_i' can be computed as

$$p(l_k|l_i') = p(l_k) \cdot p(s_{g_k}|l_{g_i}') \cdot p(k), \forall k \in N \quad (37)$$

and the adversary chooses the final inferred location $\hat{l}_i$ as

$$\hat{l}_i = \arg\max_{l_k} \; p(l_k|l_i'). \quad (38)$$

B. EXTERNAL ADVERSARY

In this case, we assume that the adversary is an untrusted third party who eavesdrops on users' request packets while they are accessing the LBS. A request packet includes the users' pseudo locations together with the timestamps, user IDs, and LBS service thresholds, which can be written as $Req(i) = \{(x_i', y_i'), t_i, \text{userID}_i, r_{\text{LBS}_i}\}$. Therefore, the malicious adversary has prior knowledge of all the request packets and the implemented LPPMs in the location service provider. With

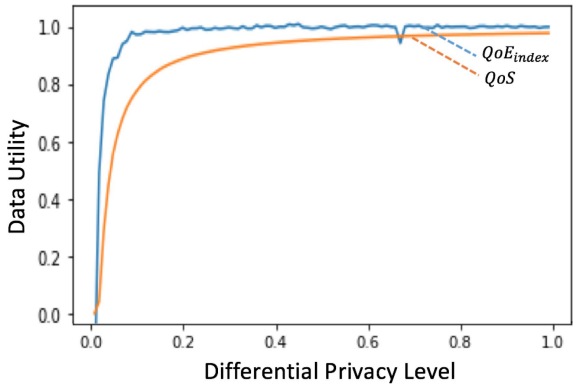


FIGURE 6. Relationship between privacy and data utility.

such assumptions, we use a HMM to characterize the adversary's actions to infer users' true locations.

HMM is a statistical Markov model and consists of five elements: hidden states $\mathbb{S}$, observable states $\mathbb{O}$, initial state probability matrix π, transition probability matrix A and emission probability matrix B [38]. In our malicious adversary model, we divide the whole service-provided area into a square grid with N regions (or tiles) and each region is a state in our HMM model. Then, we define the hidden states to be the users' true location sequences and the observable states to be the pseudo-locations. Since the user's initial position is unknown, we assume that the initial state probability matrix is an uniform distribution as

$$\pi = \left[\underbrace{\frac{1}{N}, \frac{1}{N}, \ldots, \frac{1}{N}}_{N} \right]. \tag{39}$$

The transition probability between any two hidden states can be described as $a_{ij} = \Pr(s_j|s_i)$, which denotes the probability of the next state being s_j under the condition that the current state is s_i. In our adversary model, we assume that the user's next state should be in an adjacent or the same state as the current state. Therefore, there are nine possibilities for the user's next state (except for boundary conditions) and we set the transition probability between two hidden states with the same vertex or same edge as $\frac{1}{9}$. The emission probability b_{ij} represents the probability that the hidden state s_i performs as the observable state o_j, which is more complex to compute than the transition probability. Thus, we use Monte Carlo sampling to obtain the emission probability matrix. With the above information, the malicious adversary can infer the user's true locations as the hidden states by using the Viterbi algorithm with given observable state sequences.

VII. PERFORMANCE EVALUATION

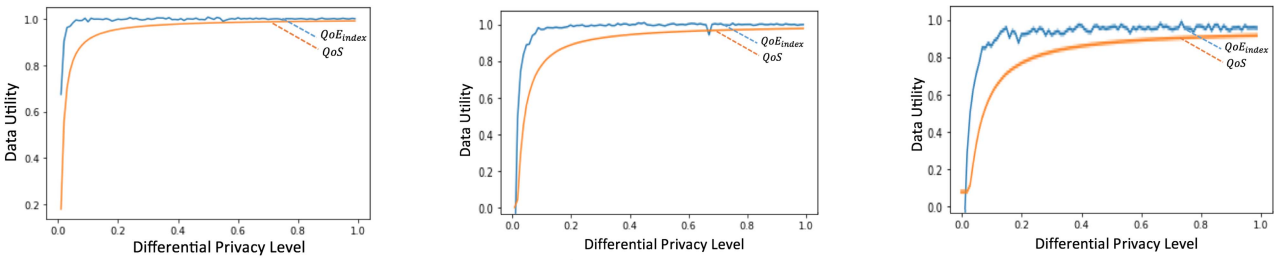
In this section, we present the results to evaluate the proposed LPPM. Fig. 6 is generated based on the Dianping dataset and shows the relationship between location privacy and data utility. In Fig. 6, two different quantification metrics of data utility are considered. The value of QoS represents the percentage of remaining useful application service after the location perturbation. QoE_{index} represents the ratio of average QoE values based on true locations and pseudo-locations fulfilling the requirement of LBS threshold.

The horizontal axis in Fig. 6 shows the differential privacy level. Intuitively, the higher the differential privacy level is, the lower the location privacy preservation level the user obtains. With an increase in the differential privacy level, QoS and QoE_{index} decrease. The value of QoE_{index} is always higher than QoS for a given differential privacy level. Thus, if we use QoE to measure the data utility, we do not need to sacrifice too much privacy to obtain the same service performance. Furthermore, this conclusion can help to find the tradeoff between location privacy and data utility. As for QoE, when the differential privacy level changes from 0 to 0.1, the user's experience quality increases greatly. However, when the differential privacy level increases from 0.3 to 1, the user's experience quality does not change significantly. There is a similar pattern for QoS, but its rate of change with ϵ is slower as compared to QoE. Therefore, we do not need to sacrifice the user's location privacy too much for higher data utility since the tradeoff level can obtain similar results, which is more cost-effective.

The effectiveness of the requirement of LBS threshold in our framework is shown in Fig. 7. In Fig. 7(a), (b) and (c), the values of r_{LBS} are set as 500 m, 200 m and 100 m, respectively. With the same privacy level, higher QoS and QoE_{index} can be obtained by using a larger r_{LBS}. We can also observe that the influence of r_{LBS} on QoS is more obvious than QoE_{index}, which implies that QoE may be more subjective and reflective of the user's real experience rather than just computing the overlapping region (as done in QoS). The results above can provide visually effective suggestions for users to select the appropriate data utility and the tradeoff levels.

We also compare the proposed LPPM with other existing location privacy preserving methods to validate our proposed framework and metrics. We compare the proposed framework with the geo-indistinguishability framework [14], which provides a mechanism to protect the user's location privacy and enhance the performance of LBSs. Fig. 8 shows the QoE loss between the geo-indistinguishability framework and our proposed framework, and the QoE performance of our system is always better than that of the geo-indistinguishability method. Since the geo-indistinguishability framework uses the notion of Area of Retrieval (AOR) rather than the Area of Interest (AOI) to provide LBSs to the user, too many unsatisfied services are reported as the service results. Although the AOI can include more services, the average QoE among all the reported services may not be good and some of the services may be too far away from the user, especially when the privacy level is small.

We evaluate the proposed adversary models and provide the privacy analysis in Fig. 9. The inference location errors are defined as the distances between inferred locations and true locations and the pseudo-location errors are defined as



(a) $r_{\mathrm{LBS}} = 500$ (b) $r_{\mathrm{LBS}} = 200$ (c) $r_{\mathrm{LBS}} = 100$

FIGURE 7. Relationship between privacy and data utility with different parameters.

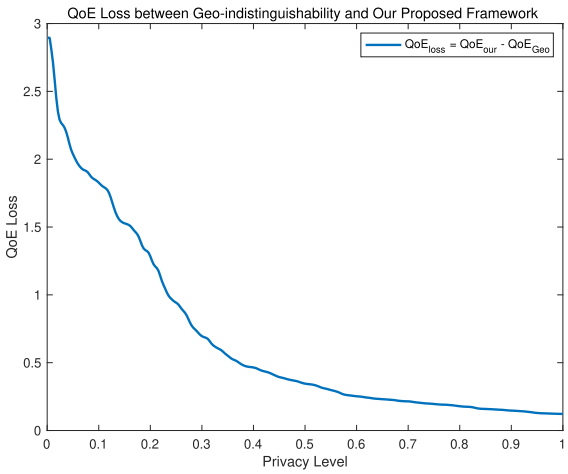


FIGURE 8. Comparison with geo-indistinguishability.

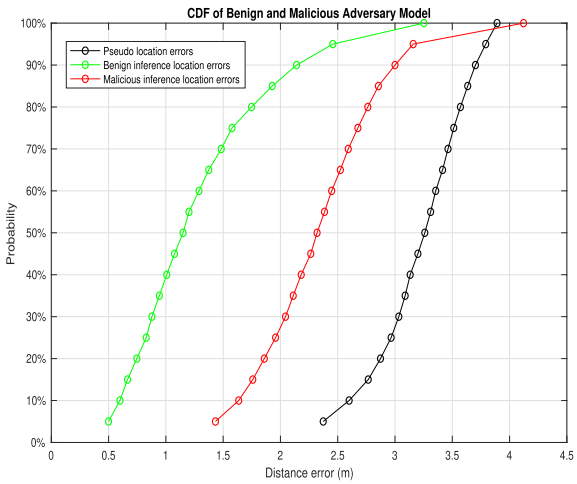


FIGURE 10. CDF of adversary model.

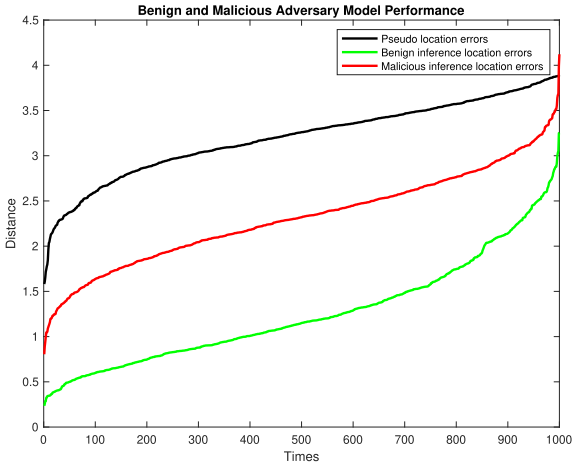


FIGURE 9. Adversary model performance.

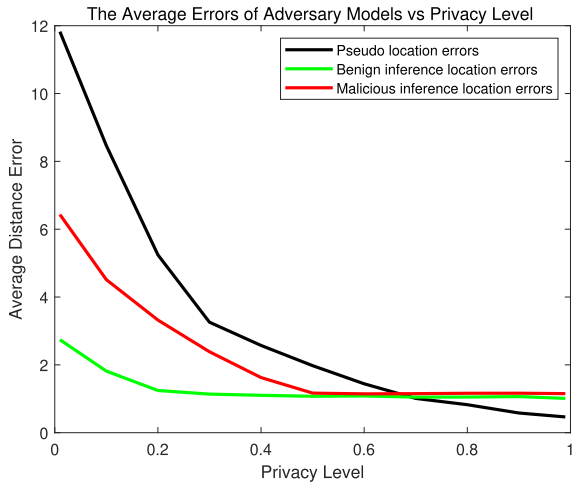


FIGURE 11. Average errors of adversary models versus privacy level.

the distances between pseudo locations and true locations. To simulate the malicious adversary, we use the random walk algorithm to generate random sequences of users' true locations. Moreover, we test the two adversary models using the same dataset of true locations and pseudo locations, which are generated by the proposed LPPM.

Fig. 9 shows that the internal adversary model can, to some extent infer users' real locations and reduce the location errors by using the inferred locations instead of using pseudo locations. In our experiment, the average pseudo location error is 3.2571 and the average inference location error for the internal adversary is 1.1380, which is almost 30% of the average pseudo location error. As for the external adversary, the accuracy of the inferred locations is lower than the internal adversary since the internal adversary has more useful prior knowledge than the external adversary. The external adversary can reduce the average location errors from 3.2571 to 2.3862.

Hence, by implementing our proposed adversary models, an internal, honest-but-curious adversary may provide better services for users than using pseudo locations directly and the external adversary can infer users' true locations to a limited extent. Even if the adversaries attempt to infer the possible real locations by implementing the adversary models, they cannot obtain the users' true locations with high accuracy. Thus, our proposed framework is robust and can still preserve users' location privacy against both the internal and external adversary models.

To show the performance of the internal and external adversary models more intuitively, the comparison of the cumulative distribution functions (CDFs) of distance errors of the proposed LPPM, internal adversary inference, and external adversary inference is shown in Fig. 10. In addition, Fig. 11 illustrates the average error comparison among the pseudo locations, the internal adversary inferred locations, and the external adversary inferred locations as a function of the privacy level. When the privacy level is small, both the internal adversary and the external adversary can reduce the average distance errors to a large extent and make the inferred location be close to the user's real locations than the pseudo locations. However, the internal and external inference errors will become even larger than the pseudo location errors when the privacy level increases. Therefore, the internal and external adversary models have the problem of over-inference when the user's location privacy requirement is not high.

VIII. CONCLUSION

This paper proposed a framework for location privacy preservation that consists of four components: localization, location privacy preservation mechanism, location-based service, and adversary models. The framework enables users to find the tradeoff between location privacy and data utility based on loc-correlated privacy. We bridged the gap between the location privacy quantification and the subjective data utility metrics and used two metrics: QoS and QoE. We also provided the privacy analysis by introducing two adversary models with different initial intentions and illustrated how the proposed LPPM protects users' location privacy against the adversarial algorithms. Finally, we evaluated the proposed system with a real dataset and analyzed the performances under different settings. The tradeoff analysis demonstrated that the proposed framework is cost-effective since we do not need to sacrifice user's location privacy superfluously for better data utility. The proposed framework is robust and adequate to protect user's location privacy.

In our future work, we will consider the location correlations to effectively preserve users' trajectory privacy. The current proposed framework needs a trusted third party to help realize the location privacy preservation. However, publishing the trajectory data or the raw sensor data used for estimating the user's trajectory to a third party may have privacy issues while the user is accessing LBS. We plan to follow the LDP principle and implement the trajectory privacy preservation mechanism locally.

REFERENCES

[1] A. R. Beresford and F. Stajano, "Location privacy in pervasive computing," *IEEE Pervasive Comput.*, vol. 2, no. 1, pp. 46–55, Jan.-Mar. 2003.

[2] B. Hoh, M. Gruteser, H. Xiong, and A. Alrabady, "Preserving privacy in GPS traces via uncertainty-aware path cloaking," in *Proc. 14th ACM Conf. Comput. Commun. Secur.*, 2007, pp. 161–171.

[3] M. Ghaffari, N. Ghadiri, M. H. Manshaei, and M. S. Lahijani, "P^4QS: A peer-to-peer privacy preserving query service for location-based mobile applications," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 9458–9469, Oct. 2017.

[4] C. A. Ardagna, M. Cremonini, S. D. C. di Vimercati, and P. Samarati, "An obfuscation-based approach for protecting location privacy," *IEEE Trans. Dependable Secure Comput.*, vol. 8, no. 1, pp. 13–27, Jan./Feb. 2011.

[5] P. Zhao, W. Liu, G. Zhang, Z. Li, and L. Wang, "Preserving privacy in WiFi localization with plausible dummy locations," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11909–11925, Oct. 2020.

[6] S. Mascetti, D. Freni, C. Bettini, X. S. Wang, and S. Jajodia, "Privacy in geo-social networks: Proximity notification with untrusted service providers and curious buddies," *Int. J. Very Large Data Bases*, vol. 20, no. 4, pp. 541–566, 2011.

[7] N. Ravi, C. M. Krishna, and I. Koren, "Enhancing vehicular anonymity in its: A new scheme for mix zones and their placement," *IEEE Trans. Veh. Technol.*, vol. 68, no. 11, pp. 10372–10381, Nov. 2019.

[8] C. Dwork, "Differential privacy," *Encyclopedia of Cryptography and Security.* Berlin, Germany: Springer, 2011, pp. 338–340.

[9] R. Shokri, "Privacy games: Optimal user-centric data obfuscation," in *Proc. Privacy Enhancing Technol.*, 2015, pp. 1–17.

[10] K. Micinski, P. Phelps, and J. Foster, "An empirical study of location truncation on Android," in *Proc. Mobile Secur. Technoligies*, 2013, pp. 1–10.

[11] W. Wu, A. Arefin, R. Rivas, K. Nahrstedt, R. Sheppard, and Z. Yang, "Quality of experience in distributed interactive multimedia environments: Toward a theoretical framework," in *Proc. 17th ACM Int. Conf. Multimedia*, 2009, pp. 481–490.

[12] R. Shokri, G. Theodorakopoulos, C. Troncoso, J.-P. Hubaux, and J.-Y. Le Boudec, "Protecting location privacy: Optimal strategy against localization attacks," in *Proc. ACM Conf. Comput. Commun. Secur.*, 2012, pp. 617–627.

[13] R. Shokri, G. Theodorakopoulos, J.-Y. Le Boudec, and J.-P. Hubaux, "Quantifying location privacy," in *Proc. IEEE Symp. Secur. Privacy*, 2011, pp. 247–262.

[14] M. E. Andrés, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, "Geo-indistinguishability: Differential privacy for location-based systems," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2012, pp. 901–914.

[15] N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, "Optimal geo-indistinguishable mechanisms for location privacy," in *Proc. ACM Conf. Comput. Commun. Secur.*, 2014, pp. 251–262.

[16] M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking," in *Proc. 1st Int. Conf. Mobile Syst. Appl. Serv.*, 2003, pp. 31–42.

[17] P. Zhao et al., "P^3-LOC: A privacy-preserving paradigm-driven framework for indoor localization," *IEEE/ACM Trans. Netw.*, vol. 26, no. 6, pp. 2856–2869, Dec. 2018.

[18] H. Kido, Y. Yanagisawa, and T. Satoh, "An anonymous communication technique using dummies for location-based services," in *Proc. 1st Int. Conf. Pervasive Serv.*, 2005, pp. 88–97.

[19] P. Shankar, V. Ganapathy, and L. Iftode, "Privately querying location-based services with sybilquery," in *Proc. 11th Int. Conf. Ubiquitous Comput.*, 2009, pp. 31–40.

[20] M. Wernke, F. Dürr, and K. Rothermel, "PShare: Ensuring location privacy in non-trusted systems through multi-secret sharing," *Pervasive Mobile Comput.*, vol. 9, no. 3, pp. 339–352, 2013.

[21] A. R. Beresford and F. Stajano, "Mix zones: User privacy in location-aware services," in *Proc. IEEE 2nd Annu. Conf. Pervasive Comput. Commun. Workshops*, 2004, pp. 127–131.

[22] T. Feng, W.-C. Wong, S. Sun, Y. Zhao, and Z. Zhang, "Location privacy preservation and location-based service quality tradeoff framework based on differential privacy," in *Proc. 16th Workshop Positioning Navigation Commun.*, 2019, pp. 1–6.

[23] K. Brunnström et al., "Qualinet white paper on definitions of quality of experience," *Eur. Netw. Qual. Experience Multimedia Syst. Serv. (COST Action IC 1003)*, White Paper, Mar. 2013.

[24] P. Gandotra, R. K. Jha, and S. Jain, "Sector-based radio resource allocation (SBRRA) algorithm for better quality of service and experience in device-to-device (D2D) communication," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 5750–5765, Jul. 2018.
[25] A. Bera, S. Misra, and C. Chatterjee, "QoE analysis in cache-enabled multi-UAV networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 6, pp. 6680–6687, Jun. 2020.
[26] R. K. Mok, E. W. Chan, and R. K. Chang, "Measuring the quality of experience of HTTP video streaming," in *Proc. IFIP/IEEE 12th Int. Symp. Integr. Netw. Manage. Workshops*, 2011, pp. 485–492.
[27] J. Joskowicz and J. C. L. Ardao, "A parametric model for perceptual video quality estimation," *Telecommun. Syst.*, vol. 49, no. 1, pp. 49–62, 2012.
[28] A. Balachandran, V. Sekar, A. Akella, S. Seshan, I. Stoica, and H. Zhang, "Developing a predictive model of quality of experience for internet video," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 43, no. 4, pp. 339–350, 2013.
[29] Y. Liu and J. Y. Lee, "A unified framework for automatic quality-of-experience optimization in mobile video streaming," in *Proc. 35th Annu. IEEE Int. Conf. Comput. Commun.*, 2016, pp. 1–9.
[30] M. Becvarik and M. Devetsikiotis, "Modeling of user quality of experience in location aware smart spaces," in *Proc. Digit. Media Ind. Academic Forum*, 2016, pp. 207–212.
[31] C. Dwork, "A firm foundation for private data analysis," *Commun. ACM*, vol. 54, no. 1, pp. 86–95, 2011.
[32] S. Xiong, A. D. Sarwate, and N. B. Mandayam, "Randomized requantization with local differential privacy," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2016, pp. 2189–2193.
[33] C. Dwork, "Differential privacy," in *Proc. 33rd Int. Conf. Automata Lang. Program.-Volume Part II*, 2006, pp. 1–12.
[34] C. Dwork et al., "The algorithmic foundations of differential privacy," *Found. Trends Theor. Comput. Sci.*, vol. 9, no. 3/4, pp. 211–407, 2014.
[35] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*. Champaign, IL, USA: Univ. Illinois Press, 1998.
[36] Y. Zhang, G. Lai, M. Zhang, Y. Zhang, Y. Liu, and S. Ma, "Explicit factor models for explainable recommendation based on phrase-level sentiment analysis," in *Proc. 37th Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2014, pp. 83–92.
[37] A. C. Cameron and F. A. G. Windmeijer, "R-squared measures for count data regression models with applications to health-care utilization," *J. Bus. Econ. Statist.*, vol. 14, no. 2, pp. 209–220, 1996.
[38] Y. Lu, D. Wei, Q. Lai, W. Li, and H. Yuan, "A context-recognition-aided PDR localization method based on the hidden Markov model," *Sensors*, vol. 16, no. 12, 2016, Art. no. 2030.

TIANYI FENG received the B.Eng. degree from Beijing University of Posts and Communications, Beijing, China, in 2017. She received her Ph.D. degree in wireless communication and network from National University of Singapore, Singapore, in 2022. Her research interests include indoor localization, privacy preservation, and wireless sensing.

ZHIXIANG ZHANG (Graduate Student Member, IEEE) received the B.Eng. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2018, and the M.Sc. degree in 2019 from the National University of Singapore, Singapore, where he is currently working toward the Ph.D. degree with the Department of Electrical and Computer Engineering. His research interests include Internet of Things, privacy preservation, and machine learning.

WAI-CHOONG (LAWRENCE) WONG (Life Senior Member, IEEE) received the B.Sc. and Ph.D. degrees in electronic and electrical engineering from Loughborough University, Loughborough, U.K., in 1976 and 1980, respectively. He was an Emeritus Professor with the Department of Electrical and Computer Engineering, National University of Singapore (NUS), Singapore. He was a member of Technical Staff with AT&T Bell Laboratories, Crawford Hill Lab, Holmdel, NJ, USA, from 1980 to 1983. From 2002 to 2006, he was the Executive Director of the Institute for Infocomm Research (I2R), Agency for Science, Technology and Research (A*STAR), Singapore. In 1983, he joined NUS, where he has been serving in various leadership positions at the department, faculty, and university levels. His research interests include wireless and sensor networks and systems, ambient intelligent platforms, localization, and source matched transmission techniques, with more than 300 publications and five patents in these areas. Dr. Wong was the recipient of several awards, including the IEE Marconi Premium Award in 1989, IEEE Millennium Award in 2000, e-nnovator Awards (Open Category) in 2000, Best Paper Award at the IEEE International Conference on Multimedia and Expo in 2006, Best Paper Award at the Second International Conference on Ambient Computing, Applications, Services, and Technology in 2012, and SupercomputingAsia HPC Network Achievement Award in 2023. He is a member of IEEE-Eta Kappa Nu.

SUMEI SUN (Fellow, IEEE) is currently a Distinguished Institute Fellow, and Acting Executive Director of Institute for Infocomm Research (I2R), Agency for Science, Technology, and Research (A*STAR), Singapore. She also holds an adjunct appointment with the National University of Singapore, and joint appointment with the Singapore Institute of Technology, both as a Full Professor. Her current research interests include next-generation wireless communications, joint communication-sensing-computing-control design, industrial Internet of Things, applied deep learning and artificial intelligence. She was the Editor-in-Chief of IEEE Open Journal of Vehicular Technology(OJVT) during 2019-2023, is the Editor-at-Large of OJVT, and a member of the IEEE Vehicular Technology Society Board of Governors during 2022–2024.

BIPLAB SIKDAR (Senior Member, IEEE) received the B.Tech. degree in electronics and communication engineering from North Eastern Hill University, Shillong, India, in 1996, the M.Tech. degree in electrical engineering from the Indian Institute of Technology, Kanpur, India, in 1998, and the Ph.D. degree in electrical engineering from the Rensselaer Polytechnic Institute, Troy, NY, USA, in 2001. He was on the faculty of Rensselaer Polytechnic Institute, Troy, NY, USA, from 2001 to 2013, first as an Assistant and then as an Associate Professor. He is currently an Professor with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore. He is currently the Head of Department with the Department of Electrical and Computer Engineering and the Director of the Cisco-NUS Corporate Research Laboratory. His research interests include wireless network, and security for IoT and cyber physical systems. Dr. Sikdar is a member of Eta Kappa Nu and Tau Beta Pi. He was an Associate Editor for the IEEE Transactions on Communications, IEEE Transactions on Mobile Computing and IEEE Internet of Things Journal, and currently serves on the editorial board of IEEE Open Journal of Vehicular Technology.