Exploiting Temporal Diversity for Privacy-Aware Location Based Services in Mobile Networks

Xiaofan He, Member, IEEE, Richeng Jin, Student Member, IEEE and Huaiyu Dai, Fellow, IEEE

Abstract—Although location-based services (LBSs) bring the unprecedented convenience to our daily life, it may lead to a severe breach of location privacy when users contact untrusted or compromised LBS providers. Despite that various privacy protection mechanisms have been developed for LBS, their performances may be fundamentally limited by the ambient physical environment. Consequently, mobile users usually experience a temporal diversity in the achievable privacy when traveling along their routes. However, to the best of our knowledge, an appropriate location privacy metric that can capture the influence from the ambient environment is still missing in the literature, and also, none of the existing location privacy protection methods can properly exploit such temporal diversity. With this consideration, new ambient environmental information based location privacy metrics are proposed in this work, together with a stochastic model that can capture the variation of these privacy metrics along the user’s route. Based on this modeling, a new optimal stopping based LBS access algorithm is developed so as to enable mobile users to fully leverage the temporal diversity and achieve a substantially better performance. The effectiveness of the proposed scheme is corroborated through both numerical results and simulations over real-world road maps.

Index Terms—Location privacy, privacy metric, location-based service (LBS), temporal diversity, optimal stopping.

I. INTRODUCTION

With the rapid development of wireless communication technology and mobile devices, location-based service (LBS) has emerged as a promising way of improving our quality of life [1–5]. In an LBS network, mobile users are allowed to send queries based on their current locations to nearby LBS providers to collect desired information. For example, a driver can use LBSs to locate the cheapest gas station in the vicinity, and a tourist can use LBSs to find a good local restaurant. Nonetheless, as each query often contains sensitive location information, without an adequate protection mechanism, users’ privacy may be severely jeopardized when contacting an untrusted or compromised LBS provider, which in turn may cause substantial financial losses and even life-threatening problems [6–8].

With this consideration, various location privacy protection mechanisms have been developed in the literature [9–17]. Spatial cloaking is one of the most widely employed techniques for privacy-aware LBS access (e.g., [9–11, 18, 19]). In this type of method, a user’s true location is blurred into a large cloaked area that covers sufficiently many other users such that the LBS provider cannot accurately pinpoint the querying user. An alternative to spatial cloaking is the mix-zone technique (e.g., [13–15]). This type of method assumes a dedicated area, called mix-zone, and when multiple users enter this zone, they are allowed to exchange their pseudonyms, so as to avoid being continuously tracked by the LBS providers. For both types of approaches, when there are insufficient neighboring users in the cloaking area or the mix zone, dummy users can be created for privacy protection [16–18].

Nevertheless, in practice, no matter what protection mechanism is in use, some fundamental limits on the privacy protection performance may be imposed by the ambient physical environment, such as street layout, building location, local user distribution, etc. For example, when a mobile user generates a dummy location at a place with extremely low user density, it can be easily identified as a falsified location by the LBS provider [20]. As another example, when a driver is driving on a rural road with grasslands on both sides, he can only create fake locations on the road, which may severely limit the location privacy protection performance, as compared to the case when he is driving in a more privacy friendly urban area with highly connected streets and roads. Consequently, the best attainable privacy protection may vary over time as the mobile user traveling along its route. This phenomenon may be treated as the temporal diversity in privacy-aware LBS access problems. Although many LBS queries (e.g., finding the cheapest gas station) are not real-time requests, most of the existing works focus on the instant LBS access problem, where it is assumed that the mobile user will access the first available LBS immediately, and hence fails to exploit such temporal diversity. Actually, as similar LBSs may be available at multiple places, for a delay tolerant query, the mobile user can wait and send out the query till he reaches a privacy-friendly location. To the best of our knowledge, this work is among the first to exploit temporal diversity to improve the performance of LBS related location privacy in mobile networks.

To formalize the idea mentioned above and derive the optimal access strategy for the mobile user, an optimal stopping based formulation of the privacy-aware LBS access problem is considered in this work. Nonetheless, several challenges prevent us from directly applying existing results. Firstly, existing location based privacy metrics can be roughly divided into k-anonymity (and its variants [21, 22]) based metrics
[9, 23], entropy based metrics [7, 19, 24], differential privacy based metrics [11, 25], and error-distance based metrics [26, 27]. However, none of these metrics can adequately reflect the influence on location privacy from the ambient environment. Considering this, several ambient information based privacy metrics are proposed in this work, in the exemplary context of vehicular network LBS.\(^1\) Secondly, to the best of our knowledge, existing privacy metrics mainly focus on static scenarios and cannot properly characterize the stochastic properties of the variation of the potential privacy loss along the user’s route. To facilitate the exploitation of the temporal diversity in the mobile user’s location privacy, a reflected random walk [31] based stochastic model is proposed in this work. In addition, inference methods are developed to acquire good estimates of the model parameters. The third challenge of exploiting temporal diversity in the privacy-aware LBS access problem is caused by a mismatch between the assumption in the classic optimal stopping problem and the practical setting. Particularly, unlike assumed in the classic optimal stopping problems [32], an LBS may not be available to the mobile user at every timeslot, and instead, the appearance of LBS providers may follow a certain random process. For this reason, a new optimal stopping based algorithm is devised in this work to find the best possible LBS access strategy for the mobile user, and the convergence property of the proposed algorithm is established as well. The contribution of this work is summarized as follows.

- New location privacy metrics that can incorporate the impact from the ambient environment on privacy are proposed. Also, a reflected random walk based stochastic model is developed to capture the spatial variation of these privacy metrics, along with the corresponding parameter estimation methods.
- A novel and provably convergent optimal stopping algorithm is developed to adequately exploit the temporal diversity for better privacy protection.
- The effectiveness of the proposed method is validated through both a numerical example and simulations over two real-world city maps.

The rest of this paper is organized as follows. Section II introduces the optimal stopping formulation of the privacy-aware LBS access problem, along with the proposed ambient environment based privacy metrics and the corresponding parameter estimation methods. The proposed optimal LBS accessing algorithm and corresponding analytic results are presented in Section III. In Section IV, the effectiveness of the proposed method is corroborated by both numerical and real-world examples. Lastly, Section V concludes the paper.

II. PROBLEM FORMULATION AND SYSTEM MODEL

In this section, the optimal stopping formulation of the privacy-aware LBS access problem is presented first. Then, the subsequent discussion is devoted to finding proper metrics and stochastic model for characterizing the user’s privacy loss when accessing an LBS. Lastly, inference methods for the associated model parameters are presented.

A. Problem Formulation

Consider a mobile user seeking for a certain type of LBS along its travel route. Each LBS provider in the region is assumed to periodically broadcast some greeting messages to nearby mobile users. Whenever the mobile user receives a greeting message, it decides whether or not to access the corresponding LBS.\(^2\) If the mobile user determines not to access the LBS, it receives nothing and continues to seek a suitable LBS. If the user determines to access an LBS appearing at time \(t\), it receives a utility \(q_0\) depending on the service type. On the other hand, whenever the mobile user accesses an LBS, the LBS provider can (roughly) pinpoint the user’s location within a circle of radius \(r\) centered at \(x_{LBS}\), where \(r\) and \(x_{LBS}\) are the communication range and the location of the LBS provider. Not only this, the LBS provider can further narrow down the range of user location by leveraging information about the ambient environment. For example, for the LBSs in vehicular networks, the LBS can further reduce its ambiguity about the user’s location based on the road structure in the vicinity, as vehicles are rarely off the roads. Apparently, this will inflict a privacy loss \(l(x_{LBS}; M)\) on the mobile user, where \(M\) represents the regional map. Therefore, the overall nominal reward to the mobile user can be modeled as

\[
r_t = \begin{cases} 
q_0 - l(x_t; M), & \text{if access,} \\
0, & \text{otherwise,}
\end{cases}
\]

(1)

where \(x_t\) represents the location of the LBS met by the mobile user at timeslot-\(t\). The actual reward to the mobile user is given by \(r_t = \alpha^t \cdot r_t\) where \(\alpha\) is a discounting function at time \(t\) with parameter \(\alpha \in [0, 1]\). Here, \(\alpha\) is used to capture the mobile user’s preference of obtaining the reward as early as possible; a smaller \(\alpha\) indicates a less patient mobile user. In addition, it is worth mentioning that this setting includes the conventional instant LBS access problems as a special case with \(\alpha = 0\) (under the convention \(0^0 = 1\)).

Naturally, the privacy-aware LBS access problem described above can be formulated as an infinite horizon Markov optimal stopping problem [32] with the associated reward at each time \(t_i\) given by \(r_i = f_i(r_i)\). However, as compared to the classic optimal stopping problems, a key difference here is that the mobile user is not guaranteed to meet an LBS provider in every timeslot and the time interval \(u_i\)’s between meeting two LBS providers is random. For this reason, a new optimal stopping algorithm has to be devised to find the optimal LBS access strategy for the mobile user. But before presenting the detailed algorithm, some more discussions are devoted to the privacy loss function \(l(x; M)\) in the next subsection.

\(^1\)This work mainly focuses on the privacy loss of one-shot LBS access and hence location privacy is sufficient. The more sophisticated trajectory privacy [17, 28, 29] for continuous LBS [30] is beyond the scope of this work.

\(^2\)It is assumed in this work that, when greeting messages from multiple LBS providers are received at the same time, the mobile user will randomly pick one among them.
B. Exemplary Privacy Metrics

As discussed previously, the privacy loss function \( l(x; M) \) should capture the fundamental impact on user’s privacy imposed by the ambient environment. To this end, three road structure based privacy metrics are proposed in this subsection, serving as concrete examples of \( l(x; M) \) in the context of LBS in vehicular networks. To the best of our knowledge, this work is among the first to study such ambient environmental information related privacy metrics.

\[ A \]

![Example 1](a) Example 1 ![Example 2](b) Example 2

![Example 3](c) Example 3 ![Example 4](d) Example 4

Fig. 1: Examples of LBS providers with different ambient environments.

Area based privacy metric \( l_A \): The essential idea of many existing location privacy protection mechanisms is to generate dummy users at locations different from the true user location. Apparently, when more candidate locations are available, better privacy protection can be achieved. In the application of vehicular networks, since candidate locations of the dummy users cannot be off the roads, a set of roads occupying a smaller area usually implies a potentially more severe privacy breach. For example, accessing the LBS in Fig. 1(a) may incur a higher privacy loss as compared to accessing the one in Fig. 1(b). With this consideration, the following area based privacy loss function \( l_A \) is proposed to evaluate the privacy loss at a given LBS provider location \( x \) in a given map \( M \)

\[
l_A(x; M) \triangleq \left( 1 - \frac{A(x, r; M)}{\pi r^2} \right) \times \omega_0, \tag{2}
\]

In the above equation, \( A(x, r; M) \) gives the total area occupied by the roads within the circle with center \( x \) and radius \( r \). The denominator \( \pi r^2 \) normalizes the range of \( l_A(x; M) \) to \([0, \omega_0]\) and \( \omega_0 \) is a weighting factor reflecting the relevant importance of privacy (as compared to \( q_0 \)).

Number of exits based privacy metric \( l_N \): After accessing an LBS, the mobile user’s future route may be predicted by the LBS provider based on the road structure in the vicinity. For example, a client of the LBS shown in Fig. 1(c) can only leave the circular area through either exit \( a \) or exit \( b \). Such information, if disclosed, may be exploited by an adversary (e.g., a robber) to intercept the victim, and thus leads to a severe financial loss and even life-threatening issues. To reflect the mobile user’s privacy concern in this aspect, a privacy loss function \( l_N \) based on the number of exits \( n \) is proposed as follows

\[
l_N(x; M) \triangleq \left( 1 - \frac{n w_{rd}}{2 \pi r} \right) \times \omega_0, \tag{3}
\]

where \( w_{rd} \) is the width of the roads and the denominator \( 2 \pi r \) ensures that \( l_N(x; M) \in [0, \omega_0] \). Under this privacy metric, the LBS in Fig. 1(b) is considered to be safer than that in Fig. 1(c), even though they produce similar values under metric \( l_A(x; M) \).

Distance based privacy metric \( l_D \): To protect its location privacy, the mobile user usually prefers to generate fake user locations as far as possible from its real location. Comparing Fig. 1(a) and Fig. 1(d), although the area based privacy metric \( l_A \) may assign similar values to these two LBSs, the LBS shown in Fig. 1(d) is intuitively more privacy-friendly because of the more dispersed road structure there. To capture the difference in privacy loss between these two scenarios, a privacy metric \( l_D \) based on the average distance between a pair of randomly selected true location \( x_{tr} \) and dummy location \( x_{dmy} \) may be defined as follows

\[
l_D(x; M) \triangleq \frac{D_{\text{max}} - \mathbb{E}\left[||x_{tr} - x_{dmy}||_2\right]}{D_{\text{max}} - D_{\text{min}}} \times \omega_0, \tag{4}
\]

where \( D_{\text{min}} \) and \( D_{\text{max}} \) are the lower and the upper bounds of \( \mathbb{E}\left[||x_{tr} - x_{dmy}||_2\right] \), respectively, over all the LBS providers in the local region. The purpose of embedding these two parameters again is to ensure that \( l_D(x; M) \in [0, \omega_0] \).

Note that the privacy metrics proposed above by no means form a complete list. Other tailor-designed privacy metric may be developed based on the user’s particular privacy concern, the underlying applications, as well as the prior information. For example, when prior information about vehicle density is available, the privacy loss function proposed in [20] may be readily adapted for vehicular network LBS applications. In addition, it is worth emphasizing that the algorithm developed in this work (presented later in Section III) is rather general and is not restricted to the particular form of any privacy metric mentioned above.

C. A Stochastic Model of \( l(x) \)

Based on the previous discussions, it is not difficult to realize that the potential privacy loss \( l(x; M) \) of accessing a nearby LBS changes over time as the mobile user travels along its route. To capture such temporal characteristic of the privacy metrics, a reflected random walk [31] based stochastic model is proposed in this subsection. For simplicity, the parameter \( M \) of \( l(x; M) \) will be omitted in the subsequent discussion, and it is assumed that the privacy loss is quantized with a quantization step size \( \epsilon > 0 \).

In the proposed stochastic model, for any two consecutive different LBS provider locations \( x_t \) and \( x_{t+\epsilon} \) along the mobile
user’s route, the corresponding privacy losses \( l(x_t) \) and \( l(x_{t'}) \) are assumed to admit
\[
l(x_{t'}) = l(x_t) + b\left(\frac{p(x_t, x_{t'})}{m_0}\right),
\]
where \( p(x_t, x_{t'}) \) is the path length from \( x_t \) to \( x_{t'} \), and \( m_0 \) is the unit length. \( b(n) \) is a reflected random walk [31] governed by \( b(0) = 0 \) and
\[
b(n + 1) = b(n) + \delta_n l_{\text{max}} - l(x_t).
\]
In (6), \( l_{\text{min}} \) and \( l_{\text{max}} \) are the minimum and the maximum possible privacy losses, respectively. The random increment \( \delta_n \)'s are assumed to be independently and identically distributed over the set \( \{0, \pm \epsilon, \ldots, \pm \gamma \epsilon\} \) with the quantization level \( \gamma \) a finite positive integer. The corresponding probability distribution of \( \delta_n \) will be denoted by \( \{p_i\}_{i=1, \ldots, 2\gamma+1} \), where
\[
p_i \triangleq \mathbb{P}(\delta_n = (i - 1 - \gamma)\epsilon), \ i = 1, \ldots, 2\gamma + 1.
\]
In practice, the parameters \( \epsilon \) and \( \gamma \) depend on the spatial variation rate of the ambient environment, and larger \( \epsilon \) and \( \gamma \) correspond to a more heterogeneous environment. In addition, the operator \( \lfloor x \rfloor \triangleq \min\{\max\{a, z\}, b\} \) ensures that the trajectory of \( l(x) \) always stays within \([l_{\text{min}}, l_{\text{max}}]\). When the mobile user moves with an average speed \( v \), (5) may be approximated by
\[
l(x_{t'}) = l(x_t) + b\left(\frac{t' - t}{\tau_0}\right),
\]
where \( \tau_0 \triangleq m_0/v \).

Remark 1: The modeling in (5) intends to capture the following observation in practice: For two locations \( x_t \) and \( x_{t'} \), when they are close to each other (corresponding to a shorter path length \( p(x_t, x_{t'}) \)), the corresponding privacy losses \( l(x_t) \) and \( l(x_{t'}) \) are often similar due to the shared part of the ambient environment. As the separation between \( x_t \) and \( x_{t'} \) increases, the corresponding \( l(x_t) \) and \( l(x_{t'}) \) may be drastically different, which can be captured by adding more independent increment \( \delta_n \)'s in (6). In addition, it is not difficult to realize that all the privacy metrics proposed in Section II-B are consistent to the model given by (5) and (6) (with \( l_{\text{min}} = 0 \) and \( l_{\text{max}} = \omega_0 \)). Those privacy metrics are functions of the local road structure, and for two nearby locations \( x \) and \( x' \) along the user’s traveling path, the corresponding local road structures are likely to be similar and, hence, so are the values of the corresponding privacy metrics; and vice versa.

In addition, in the proposed stochastic model, it is assumed that the appearance of LBS providers along the user’s travel route follows a Poisson process. More specifically, let \( t_i \) be the time that the mobile user receives the greeting message from the \( i \)th LBS provider. Then, the lengths of the time interval \( u_i \triangleq t_{i+1} - t_i \) (for \( i = 1, 2, \ldots \)) follow independent exponential distributions with parameter \( \lambda \) (which may depend on scenario-specific factors, such as the user’s speed and the density of the LBS providers in the region).

This is an approximation since the positions of the mobile user may be different from those of the LBSs up to \( r \).

D. Parameter Extraction

In this subsection, an example based on the real map of Concord (a city in Massachusetts, USA) shown in Fig. 2 will be used to illustrate how to extract relevant parameters \( \gamma, \epsilon \) and \( p_i \)'s of the proposed stochastic model.

Fig. 2: Roads in part of Concord, Massachusetts [33].

Fig. 3: Extracted roads in Concord and generated LBS provider locations.

In particular, road information is first extracted from the map of Concord as shown in Fig. 3, and over which \( n_{\text{route}} = 1 \times 10^3 \) random routes are generated.\(^4\) Then, \( n_s \) sample locations \{\( x_i \)\}_{i=1, 2, \ldots, n_s} \) along each route are chosen with consecutive locations separated by a unit distance \( m_0 \) and then the corresponding privacy losses \( l(x_i) \)'s are computed based on the discussion in Section II-B. Considering the model specified by (5) and (6), one can realize that a difference \( \delta = l(x_{i+1}) - l(x_i) \) between the privacy losses at two consecutive sample locations can be treated as a sample of the independent increment \( \delta_n \). The histogram of the differences \( \delta \)'s collected over all the \( n_{\text{route}} \) routes are presented in Fig. 4(a), Fig. 4(b), and Fig. 4(c), respectively, for the three privacy

\(^4\)Note that these random routes are generated to facilitate parameter estimation and are not related to the actual routes of the mobile users.
metrics proposed in Section II-B. These histograms will be used to approximate the distribution of $\delta_n$. Specifically, for a given quantization level $\gamma$, a reasonable estimate of $\epsilon$ can be constructed as follows

$$\hat{\epsilon} = \frac{\max\{\hat{\delta}_{\text{max}}, \hat{\delta}_{\text{min}}\}}{2 + \gamma},$$

(9)

where $\hat{\delta}_{\text{min}}$ and $\hat{\delta}_{\text{max}}$ are the minimum and the maximum privacy losses, respectively, among the collected samples $\delta_n$. Accordingly, a good estimate of the probability $p_i$ (for $i = 1, \ldots, 2\gamma + 1$) defined in (7) could be

$$\hat{p}_i = \frac{n_i[(i-1)\epsilon - \gamma \epsilon - \hat{\delta}_i, (i-1)\epsilon - \gamma \epsilon + \frac{1}{2}\epsilon]}{n_s \cdot n_{\text{route}}},$$

(10)

where $n_i[(i-1)\epsilon - \gamma \epsilon - \hat{\delta}_i, (i-1)\epsilon - \gamma \epsilon + \frac{1}{2}\epsilon]$ denotes the number of samples falling into the range $[(i - 1)\epsilon - \gamma \epsilon - \hat{\delta}_i, (i - 1)\epsilon - \gamma \epsilon + \frac{1}{2}\epsilon]$, with ties broken equally; the denominator $n_s \cdot n_{\text{route}}$ is the total number of samples.

![Histograms](image)

Fig. 4: Histograms based on the collected data.

The LBS meeting time intervals $u_i$’s depend on not only the roads but also the LBS provider locations. To illustrate the estimation method of $\lambda$, $n_{\text{LBS}} = 100$ LBS providers are randomly generated in the region of interest and their locations are represented by the red stars in Fig. 3. Then, samples of $u_i$ can be collected based on the generated $n_{\text{route}}$ random routes and the LBS provider locations. The corresponding histogram is presented in Fig. 4(d). It can be seen from Fig. 4(d) that it is indeed reasonable to approximate the actual distribution of $u_i$ by an exponential distribution. For a given set of $n_u$ samples $\{u_j\}_{j=1}^{n_u}$ of $u_i$, the log-likelihood function is given by

$$L(\{u_j\}_{j=1}^{n_u}; \lambda) = n_u \ln(\lambda) - \lambda \sum_{j=1}^{n_u} u_j,$$

(11)

As a result, the maximum likelihood estimate of $\lambda$ can be found by applying the first order condition of optimality on (11), which yields

$$\hat{\lambda} = \frac{n_u}{\sum_{j=1}^{n_u} \hat{u}_j}.$$

(12)

Note that, although the above modeling and estimation involve certain heuristics, their effectiveness will be evidenced by the simulation results shown in Section IV. In practice, the administrator of the local region (e.g., the city’s government) can use the above method to pre-compute these parameters and send them to passing mobile users as a greeting message. Also, each LBS provider can compute its own $l(x_{\text{LBS}})$ and broadcast this information to mobile users within its communication range. The overall pre-processing procedure is summarized in Procedure 1.

### III. THE PROPOSED METHOD

With the model presented in the previous section, an optimal stopping based privacy-aware LBS access method that allows the mobile user to fully exploit the temporal diversity in privacy protection is developed in this section.

In the considered LBS access problem, the objective of the mobile user is to find an optimal access strategy $T^*_i$ that can maximize its expected actual reward $E \left[ f_{T^*_i} \left( r_{T^*_i} \right) \right]$ when meeting the $i$th LBS provider at time $t_i$ and observing a nominal reward $r_{t_i} = r$, for any $r \in \mathbb{R}$ and $i > 0$. Mathematically, $T^*_i$ is known as the optimal stopping time. More specifically, consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and the natural filtration $\{\mathcal{F}_i; i = 1, 2, \ldots\}$ with $\mathcal{F}_i \triangleq \sigma(\{r_{t_j}\}_{j=1}^{i})$. A stopping time $T$ is an (extended) random variable taking values in the set $N = \{1, 2, \ldots\}$ such that the set of events $\{\omega \in \Omega | T(\omega) \leq i\}$ is measurable with respect to the $\sigma$-algebra $\mathcal{F}_i$ for all $i$ [32]. In addition, denote by $T_k$ a stopping time that stops (i.e., ‘accesses’ in the context of privacy-aware LBS) no earlier than the appearance of the $k$th LBS provider (i.e., $t_{T_k} \geq t_k$), and let $T_k$ be the set of all possible $T_k$’s. Then, the optimal stopping time, conditioned on $r_{t_i} = r$, can be expressed as

$$T^*_i \triangleq \arg \sup_{T \in T_i} E \left[ f_{T} \left( r_{T} \right) \right] | r_{t_i} = r,$$

(13)
and accordingly, a value function \( V_i(r) \) is defined as follows to represent the corresponding optimal expected actual reward

\[
V_i(r) \triangleq \mathbb{E} \left[ f_{t_{t_i}} \left( r_{t_{t_i}} \right) \mid r_{t_i} = r \right], i = 1, 2, ..., \text{ and } \forall r \in \mathcal{R},
\]

where \( \mathcal{R} \) denotes the set of all possible nominal rewards \( r_t \).

Since the nominal reward sequence \( \{r_t\}_{t \geq 1} \) follows a time-homogeneous Markov chain determined by (1) and (8), it is not difficult to realize that \( T_i^* \) and \( V_i(r) \) defined above admit the following homogeneous properties. In particular, it has

\[
T_i^* - T_i^* = t_i - t_1, \forall i = 1, 2, ..., (15)
\]

and

\[
V_i(r) = \mathbb{E} \left[ f_{t_{t_i}} \left( r_{t_{t_i}} \right) \mid r_{t_i} = r \right] = \alpha_1 \cdot \mathbb{E} \left[ f_{t_{t_i} - t_1} \left( r_{t_{t_i} - t_1} \right) \mid r_0 = r \right] = \alpha_1 \cdot V_i (r), \quad (16)
\]

where the second equality is due to the definition of \( f \) and time-homogeneity of \( \{r_t\}_{t \geq 1} \) as well as \( t_1 \overset{d}{=} 0 \); the last equality is due to the definition of \( V_i(r) \).

Once the mobile user receives the greeting message from the \( i \)th LBS at time \( t_i \) and evaluates the nominal reward \( r_{t_i} = r \), it has to choose between two options: (1) accessing the LBS and receiving a reward \( f_{t_i}(r) \), or (2) discarding it and continuing to search for a better future reward \( V_{i+1}(r_{t_{i+1}}) \). Therefore, the following recursive equation holds

\[
V_i(r) = \max \left\{ f_i(r), \mathbb{E}_{t_{i+1}, r_{t_{i+1}}} \left[ V_{i+1}(r_{t_{i+1}}) \mid t_i, r_{t_i} = r \right] \right\},
\]

which, due to (16), is equivalent to

\[
V_1(r) = \max \left\{ r, \mathbb{E}_{u,s} \left[ \alpha^u \cdot V_1(s) \mid r_0 = r \right] \right\}. \quad (17)
\]

In the above equation, based on the stochastic model presented in Section II-C, the time interval \( u = t_{i+1} - t_i \) follows an exponential distribution \( \exp(\lambda) \) and the future nominal reward admits \( s = r + b(\frac{u}{\tau_0}) \). It is not difficult to realize that, if \( \mathbb{E}_{u,s} \left[ \alpha^u \cdot V_1(s) \mid r_0 = r \right] \) is known for all \( r \in \mathcal{R} \), the optimal LBS access decision \( d \) for the mobile user when meeting the \( i \)th LBS provider is

\[
d = \begin{cases} \text{access, if } r_{t_i} \geq \mathbb{E}_{u,s} \left[ \alpha^u \cdot V_1(s) \mid r_0 = r_{t_i} \right], \\ \text{continue, otherwise.} \end{cases} \quad (18)
\]

To evaluate \( \mathbb{E}_{u,s} \left[ \alpha^u \cdot V_1(s) \mid r_0 = r \right] \), define an \( n_r \)-dimensional vector \( g \) by stacking all the possible nominal rewards together \( g = [g_0 - l_{\text{max}}, g_0 - l_{\text{max}} + c, g_0 - l_{\text{max}} + 2c, ..., g_0 - l_{\text{min}}]^{T} \), where \( n_r = |\mathcal{R}| \) denotes the number of different nominal rewards. Further define a value vector \( V_1 = [V_1(g_0 - l_{\text{max}}), V_1(g_0 - l_{\text{max}} + c), V_1(g_0 - l_{\text{max}} + 2c), ..., V_1(g_0 - l_{\text{min}})]^{T} \). Also, let \( P_{t,k,j} \triangleq \mathbb{P} \left( r_{t_{n+1}} = g_j \mid r_n = g_i, \frac{u_n}{\tau_0} = k \right) \) be the probability that the nominal reward \( r_{t_{n+1}} \) equals \( g_j \) when encountering the \( (n+1) \)th LBS, conditioned on that \( r_n = g_i \) and that the time interval \( u_n \) between meeting the \( n \)th and the \( (n+1) \)th LBS providers admits \( k\tau_0 \leq u_n < (k+1)\tau_0 \). Due to (6) and (8), it can be verified that, for \( k \geq 0 \), \( P_{t,k,j} \) admits the following iterative relation

\[
P_{t,k+1,j} = \begin{cases} \min \{n_r, j+\gamma \} \left( \sum_{l=\max \{1, j-\gamma \}}^{1+\gamma} \sum_{s=1+\gamma+j-l}^{s=1+\gamma+j-l} p_s \right) \cdot P_{t,k,l}, \quad j = n_r, \quad (20) \\
\sum_{l=\max \{1, j-\gamma \}}^{1+\gamma+j-l} \sum_{s=1}^{n_r} p_s \cdot P_{t,k,l}, \quad 1 < j < n_r, \quad (20) \\
\sum_{j=1}^{n_r} Q_{i,j} V_i(g_j), \quad (21)
\end{cases}
\]

where \( Q_{i,j} \triangleq \frac{\lambda}{\lambda - \ln \alpha} \sum_{k=0}^{\infty} P_{t,k,j} \left( e^{(\ln \alpha - \lambda)k\tau_0} - e^{(\ln \alpha - \lambda)(k+1)\tau_0} \right) \).

By defining an \( n_r \times n_r \) matrix \( Q = [Q_{i,j}] \) and combining (17) and (21), it can be seen that the value vector \( V_1 \) admits the following recursive equation

\[
V_1 = \max \left\{ g, Q V_1 \right\}. \quad (23)
\]

**Proposition 1:** There is a unique vector \( V_1^* \) that admits (23). In addition, starting from any initial value \( V_1^{(0)} \), the sequence \( \{V_1^{(n)}\}_{n \geq 0} \) generated by the iteration

\[
V_1^{(n+1)} = \max \left\{ g, Q V_1^{(n)} \right\}, \quad (24)
\]

converges to \( V_1^* \).

**Proof:** Please see Appendix A.

Since there is no closed-form expression for \( P_{t,k,j} \), in practice, exact computation of \( Q_{i,j} \) using (22) is not feasible. Consider the following approximation that evaluates (22) up to a finite order \( k_m \):

\[
Q_{i,j}^{(k_m)} = \frac{\lambda}{\lambda - \ln \alpha} \sum_{k=0}^{k_m} P_{t,k,j} \left( e^{(\ln \alpha - \lambda)k\tau_0} - e^{(\ln \alpha - \lambda)(k+1)\tau_0} \right), \quad (25)
\]
The following proposition ensures that the above approximation has negligible influence on the evaluation of \( V_1 \) when \( k_m \) is sufficiently large.

**Proposition 2:** Let \( V \) and \( \hat{V}^{(k_m)} \) be the fixed points of the mappings \( Mz = \max\{g, Qz\} \) and \( M^{(k_m)}z = \max\{g, \hat{Q}^{(k_m)}z\} \), respectively. Then,

\[
\lim_{k_m \to \infty} ||V - \hat{V}^{(k_m)}||_\infty = 0. 
\]

**Proof:** Please see Appendix B.

The overall optimal privacy-aware LBS access algorithm implemented by the mobile user is summarized in Algorithm 1.

**Algorithm 1 Optimal Privacy-Aware LBS Access**

1. Determine a proper privacy metric \( l \in L = \{l_A, l_N, l_D\} \)
2. Enter a region and receive from the administrator a greeting message \( msg = [\gamma, \epsilon, \{\hat{p}_i\}, \lambda] \)
3. Iteratively compute \( P_{z,k,j} \) by (20)
4. Compute \( \{Q_{z,j}\} \) (approximately) by (25)
5. Iteratively compute \( V^*_i \) by (24)
6. Compute \( \mathbb{E}_{n,s} [\alpha^u \cdot V^*_i(s)|r_0 = g_1] \) by (21)
7. Set \( d = \text{continue} \)
8. while \( d = \text{continue} \) do
9. if get a greeting message from LBS provider \( i \) then
10. Decode \( \langle x_i \rangle \) from the message
11. Compute the reward \( r_t \) by (1)
12. Make a decision \( d \) according to (18)
13. if \( d = \text{access} \) then
14. Access the LBS and terminates the algorithm
15. end if
16. end if
17. Keep traveling
18. end while

The performance of the temporal diversity based privacy-aware LBS access strategy is investigated in this section. Particularly, a numerical example is considered first. The objectives are to demonstrate the performance gain brought by the proposed framework as well as to reveal some insights about the impacts of different model parameters. Then, the performance of the proposed framework on more practical scenarios are examined based on two real-world maps, the Concord city map and the Boston city map.

**A. Numerical Example**

In the considered numerical example, the number of nominal rewards is set to \( n_r = 5 \) and \( q_0 \) is set to 6. \( l_{\min} \) and \( l_{\max} \) are set to 1 and 5, respectively. In addition, \( \epsilon \) is set to 1, and \( p_1, p_2 \) and \( p_3 \) are all set to \( \frac{4}{3} \). Based these parameters, the privacy loss corresponding to each LBS provider is generated according to (8). Also, in view of Proposition 2, \( k_m \) in Algorithm 1 is set to 200, which is considered to be sufficiently large for the application of interest. The performance of the proposed method \( \mathbb{E}\left[f_{T^{*}_1}\left(r_{T^{*}_1}\right)\right] \) and that of a baseline method \( \mathbb{E}\left[r_{T_1}\right] \), in which the mobile user always accesses the first available LBS, are compared. More specifically, the relative performance gain, defined as

\[
\eta \triangleq \frac{\mathbb{E}\left[f_{T^{*}_1}\left(r_{T^{*}_1}\right)\right] - \mathbb{E}\left[r_{T_1}\right]}{\mathbb{E}\left[r_{T_1}\right]} \times 100\%,
\]

is the metric of interest here.

**Fig. 5:** Relative performance gain of the proposed method \((n_s = 5)\).

**Fig. 5** shows the performance comparison under different combinations of \( \lambda \) and \( \tau_0 \), where all the curves are obtained by averaging over \( 1 \times 10^3 \) Monte Carlo runs. First, it can be observed that the proposed method always achieves a better performance than the baseline approach. For example, when \( \alpha = 0.99 \) (corresponding to \( \ln(\frac{1}{1-\alpha}) = 4.6 \) in Fig. 5), \( \lambda = 1 \), and \( \tau_0 = 0.1 \), the proposed method can help the mobile user achieve a more than 55% extra performance gain. In addition, as the mobile user becomes more patient (corresponding to a larger \( \alpha \)), the relative performance gain increases. This indicates that the proposed algorithm can successfully enable the user to exploit the temporal diversity and thus achieve a substantially less privacy leakage. Also, it can be seen from Fig. 5 that, for a fixed \( \lambda = 1 \), the performance gain increases as \( \tau_0 \) decreases. For example, when \( \tau_0 \) is reduced from 1 to 0.1, the relative performance gain \( \eta \) increases from around 30% to nearly 60% for a mobile user with \( \alpha = 0.99 \). The reason is that a smaller \( \tau_0 \) corresponds to a faster change in the ambient environment along the user’s travel route (c.f. (8)), which provides the user a richer temporal diversity to exploit. Moreover, for a fixed \( \tau_0 = 1 \), a better performance can be observed for the smaller \( \lambda \). The is because, a smaller \( \lambda \) implies a shorter average waiting time \( u_i \) between the appearances of LBS providers and hence a better performance.

**B. Case Studies Based on Real-world Maps**

After exploring the numerical example above, two more practical scenarios are examined in the sequel and the three privacy metrics proposed in Section II-B will be considered.
The performance of the proposed method on the Concord city map (c.f. Fig. 2) is examined first. To this end, $n_{\text{route}} = 1 \times 10^5$ random routes are generated on the Concord city map along with $n_{LBS} = 100$ randomly deployed LBS providers (c.f. Fig. 3). The communication range of the LBS providers is set to $r = 5$ (unit) and $\tau_0$ is set to 1 (i.e., the mobile user’s speed is 1 unit distance per timeslot). In addition, both $\gamma_0$ and $\omega_0$ are set to 1. By setting the quantization level to $\gamma = 2$ and executing the parameter extraction procedures presented in Section II-D, the obtained estimates of the parameters are presented in Table I. Based on these estimates, the performance of the proposed scheme is shown in Fig. 6. It can be seen from Fig. 6 that the proposed algorithm outperforms the baseline approach and the corresponding relative performance gain $\eta$ increases when the mobile user is more patient (corresponding to larger $\alpha$), for all the three privacy metrics considered here. For example, when the discounting factor $\alpha$ increases from 0.999 (corresponding to $\ln \left( \frac{1}{\tau-\alpha} \right) = 6.9$) to 0.999999 (corresponding to $\ln \left( \frac{1}{\tau-\alpha} \right) = 11.5$), the relative performance gains corresponding to the three proposed privacy metrics are boosted from about 10%, 25% and 15% to about 35%, 95% and 105%, respectively. This also indicates that the proposed method is very general and remains effective under different privacy metrics. In addition, it is worth mentioning that, here, there is no intention to compare the performances across different privacy metrics; in practice, choosing the most suitable privacy metric is subject to the specific application and the user’s requirement.

<table>
<thead>
<tr>
<th>$l_4$</th>
<th>$p_1$</th>
<th>$p_2$</th>
<th>$p_3$</th>
<th>$p_4$</th>
<th>$p_5$</th>
<th>$\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.08</td>
<td>0.002</td>
<td>0.099</td>
<td>0.798</td>
<td>0.099</td>
<td>0.002</td>
<td>0.013</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$l_N$</th>
<th>$p_1$</th>
<th>$p_2$</th>
<th>$p_3$</th>
<th>$p_4$</th>
<th>$p_5$</th>
<th>$\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.27</td>
<td>0.001</td>
<td>0.041</td>
<td>0.916</td>
<td>0.041</td>
<td>0.001</td>
<td>0.008</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$l_D$</th>
<th>$p_1$</th>
<th>$p_2$</th>
<th>$p_3$</th>
<th>$p_4$</th>
<th>$p_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.23</td>
<td>0.001</td>
<td>0.033</td>
<td>0.932</td>
<td>0.043</td>
<td>0.001</td>
</tr>
</tbody>
</table>

TABLE I: Estimated parameters (Concord)

TABLE II: Estimated parameters (Boston)

To further justify the effectiveness of the proposed method, another practical scenario concerned with the Boston city map is examined. Similar to the previous case, the part of the Boston city map shown in Fig. 7 is converted into the road map presented in Fig. 8. Again, it is assumed that $n_{LBS} = 100$ LBS providers are randomly deployed in this area. After generating $n_{\text{route}} = 1 \times 10^3$ random routes on the road map and executing the parameter extraction methods, one can obtain estimates of the model parameters shown in Table II and then plug them into Algorithm 1. The performance of the proposed method for the three different privacy metrics is presented in Fig. 9. Based on Fig. 9, similar observations can be made as in the previous Concord case. This evidences the effectiveness of the proposed method across different geographical situations in practice.

Fig. 7: Roads in part of Boston, Massachusetts [34].

V. CONCLUSIONS

In this work, three new location privacy metrics that can capture the influence on privacy from the ambient environment are proposed. In addition, a stochastic model based on reflected random walk is developed to characterize the spatial variation of the location privacy along the user’s route. Based on this modeling, a new optimal stopping based privacy-aware LBS access algorithm is developed. The proposed algorithm enables the mobile users to fully exploit the temporal diversity to select the most suitable LBS provider. Corresponding analysis shows that the optimal stopping decision and values of the privacy-aware LBS access problem can be obtained through iterated computations. Results of both numerical and real-world examples show that the proposed method can achieve
If \( \|Qz\| < g_i \), then
\[
\|\|Mz - [Mz']\|\| \leq \|\|Qz - [Qz']\|\|,
\]
(30)

Therefore, the following inequality holds
\[
\|Mz - Mz'\|_{\infty} \leq \|Qz - Qz'\|_{\infty}
\]
\[
\leq \|Q\|_{\infty} \cdot \|z - z'\|_{\infty},
\]
(31)

where \( \|Q\|_{\infty} \) is the induced infinity norm of matrix \( Q \). Further notice that
\[
\|Q\|_{\infty} = \max_{1 \leq i \leq n_r} \sum_{j=1}^{n_r} |q_{i,j}|
\]
\[
= \max_{1 \leq i \leq n_r} \frac{\lambda}{\lambda - \ln \alpha} \sum_{k=0}^{\infty} \sum_{j=1}^{n_r} P_{k,j} \left( e^{(\ln \alpha - \lambda)k\tau_0} - e^{(\ln \alpha - \lambda)(k+1)\tau_0} \right) = \frac{\lambda}{\lambda - \ln \alpha} < 1, \quad (32)
\]
where the last inequality is due to \( \alpha \in [0, 1) \). The contraction property of \( M \) follows readily from (31) and (32), which completes the proof. 

**APPENDIX B**

PROOF OF PROPOSITION 2

Proposition 2 can be proved as follows.

**Proof:** Invoking similar arguments as in the proof of Proposition 1 yields
\[
\|V - \hat{V}^{(k_m)}\|_{\infty} \leq \|QV - \hat{Q}^{(k_m)}\hat{V}^{(k_m)}\|_{\infty}
\]
\[
\leq \|\hat{Q}^{(k_m)}\|_{\infty} \cdot \|V - \hat{V}^{(k_m)}\|_{\infty} + \|Q - \hat{Q}^{(k_m)}\|_{\infty} \cdot \|V\|_{\infty},
\]
which implies
\[
\|V - \hat{V}^{(k_m)}\|_{\infty} \leq \frac{\|V\|_{\infty}}{1 - \|\hat{Q}^{(k_m)}\|_{\infty}} \|Q - \hat{Q}^{(k_m)}\|_{\infty}
\]
\[
\leq \frac{\|V\|_{\infty}}{1 - \|Q\|_{\infty}} \|Q - \hat{Q}^{(k_m)}\|_{\infty}, \quad (33)
\]
where the second inequality is due to \( 0 \leq \hat{Q}_{i,j}^{(k_m)} \leq Q_{i,j}, \forall i, j \). Consequently, Proposition 2 follows readily from \( \lim_{k_m \to \infty} \|Q - \hat{Q}^{(k_m)}\|_{\infty} = 0 \). 

**REFERENCES**


[34] https://www.google.com/maps/@42.360365,-71.0583815,15.5z.