Achieving Perfect Location Privacy in Markov Models Using Anonymization

Zarrin Montazeri
Electrical and Computer Engineering Department
University of Massachusetts
Amherst, Massachusetts
Email: seyedehzarin@umass.edu

Amir Houmansadr
College of Information and Computer Sciences
University of Massachusetts
Amherst, Massachusetts
Email: amir@cs.umass.edu

Hossein Pishro-Nik
Electrical and Computer Engineering Department
University of Massachusetts
Amherst, Massachusetts
Email: pishro@engin.umass.edu

***Abstract*—Despite the popular advantages of location-based services (LBS), e.g., mobile navigation and recommender systems, they impose significant privacy threats to their users because of their mainly unrestricted access to users' raw location information over time. To limit the extent of such privacy leakage, various location privacy protection mechanisms (LPPM) have been designed, which perturb users' location information or identities before being presented to an LBS system. In this work, we extend our recent preliminary work [1] by using Markov chains to model real-world scenarios, in which a user's movements are dependent over the time domain. We define the notion of perfect location privacy, and demonstrate the feasibility of achieving the defined perfect location privacy by practical LPPM systems. We model each user's path using Markov chains and show that perfect location privacy is achievable for a user if the pseudonym of the user is changed before $O(N^{\frac{2}{|E|-r}})$ observations by the adversary, in which $|E|$ is the number of edges in the Markov model and r is the number of all possible locations. Results of our simulations are consistent with what we prove in this paper.**

I. Introduction

Today's mobile devices offer a wide range of services to their owners by recording and processing the geographic locations of their users. Such services, broadly known as location-based services (LBS), include navigation, ride-sharing, dining recommendation and auto collision warning. While such LBS applications offer a wide range of popular and important services to their users unfortunately, they impose significant privacy threats to their users because of their unrestricted access to the location information of mobile devices. Such privacy compromises can be launched by various types of adversaries including third-party applications, nearby mobile users and cellular service providers.

To protect the location privacy of LBS users various mechanisms have been designed [2]–[4], which are known as location privacy protection mechanisms (LPPM). These mechanisms tend to perturb a mobile user's information before it is disclosed to the LBS application. LPPMs that perturb location information of the users are known as *location perturbation LPPMs* and those that perturb the user's identity information are known as *identity perturbation LPPMs*. The improved location privacy by these LPPM mechanisms usually comes at the price of performance degradation for the underlying LBS systems.

In this paper, we propose a mathematical framework for location privacy for LBS services using information theory. This work is built on our preliminary work [1], in which we defined information theoretic notions of location privacy. While our preliminary study assumed each user's location information to be independent over the time domain to simplify the derivations, in this work we extend the threat model to a more complex, more realistic setting by considering dependencies between locations over time. We use Markov chains to model each user's mobility over time. Also we assume the strongest model for the adversary, i.e., we assume that the adversary has complete statistical knowledge of the users' movements. We formulate a user's location privacy based on the mutual information between the adversary's observations and the user's actual location information. We define the notion of *perfect location privacy* and show that properly designed LPPM mechanisms can achieve the defined perfect location privacy. Due to space limitations, when proving the results, we refer to [1] when applicable, and only focus on the parts that are novel and different from those we already proved in [1].

II. Related Work

Location privacy has been an active field of research over the past decade [5]–[10]. Studies in this field can be classified into two main classes: 1) works on designing effective LPPM mechanisms for specific LBS systems and platforms, and 2) works on deriving theoretical models for location privacy, e.g., by deriving metrics to quantify location privacy.

The LPPM designs can be further classified into two classes: location perturbation LPPMs and identity perturbation LPPMs. Location perturbation LPPMs aim at obfuscating the location information of the users over the time and geographical domains. Example techniques are location cloaking [5], [6] and dummy locations [7], [8].

On the other hand, identity perturbation LPPMs try to obfuscate the users' identities while using an LBS system. A common approach is called k-anonymity LPPM [3], [11] in which each LBS user's identity is kept indistinguishable within a group of k LBS users. Another approach to identity pertur-

bation LPPM is to exchange user identities (or pseudonyms) inside specified areas, called mixed-zones, as users traverse such areas in order to confuse the adversaries [12], [13].

Several works aim at quantifying location privacy protection. Shokri et al. [9], [10] define the expected estimation error of the adversary as a metric to evaluate LPPM mechanisms. On the other hand, Ma et al. [14] use uncertainty about users' location information to quantify users' location privacy in vehicular networks. Li et al. [15] define metrics to quantify the tradeoff between privacy and utility of LPPM systems. Shokri et al. [2] design LPPM mechanisms that will defeat localization attacks.

The mutual information has been used as a measure of privacy, [16], [17] but in this paper we specifically use the mutual information for location privacy. This paper aims at establishing an information theoretic framework for location privacy. Such a framework will allow us to achieve provable location privacy.

III. Framework

A. Defining Location Privacy

As we defined in [1], we consider a network with N users which uses an LPPM to support its users' location privacy. An adversary $\mathcal{A}$ is interested in identifying each user based on their location and movements. We consider this adversary to be the strongest adversary that has complete statistical knowledge of the users' movements based on the previous observations or other resources. The adversary has a model that describes users' movements as a random process on the corresponding geographic area.

Starting at time zero, users start moving. Let $X_i(t)$ be the location of user i at time t. The adversary is interested in knowing $X_i(t)$ for $i = 1, 2, ..., N$ based on her past observations through time. The adversary's observations are anonymized versions of the $X_i(t)$'s produced by the LPPM. Let $\mathbf{Y}$ be a collection of observations available to the adversary. We define *perfect location privacy* as follows:

Definition 1. User i has perfect location privacy at time t with respect to adversary $\mathcal{A}$, if and only if

$$\lim_{N\to\infty} I(X_i(t); \mathbf{Y}^M) = 0,$$

where $I(.)$ shows the mutual information and M is the number of previous observations of the adversary.

Above definitions shows that over time, the adversary's observations does not give any information about the user's location. With the assumption of $N \to \infty$, this is valid for all the applications that we consider.

Throughout this paper, in order to achieve location privacy, we use only anonymization techniques to confuse the adversary. We perform a random permutation $\Pi^{(N)}$, chosen uniformly at random among all $N!$, on the set of N users, and then assign the pseudonym $\Pi^{(N)}(i)$ to user i.

$$\Pi^{(N)} : \{1, 2, \cdots, N\} \to \{1, 2, \cdots, N\}$$

Throughout the paper, for simplicity of notations we sometimes drop the superscripts, e.g., $\Pi^{(N)} = \Pi$.

For $i = 1, 2, \cdots, N$ and time $t = 1, 2, \cdots, M$, let $\mathbf{X}_i^{(M)} = (X_i(1), X_i(2), \cdots, X_i(M))^T$ be a vector which shows the i^{th} user's locations at any time t. Using the permutation function $\Pi^{(N)}$, the adversary observes a permutation of users' location vectors, $\mathbf{X}_i^{(M)}$'s. In other words, the adversary observes

$$\begin{aligned}\mathbf{Y}^M &= \text{Perm}(\mathbf{X}_1^M, \mathbf{X}_2^M, \cdots, \mathbf{X}_N^M; \Pi^{(N)}) \\ &= (\mathbf{X}^M_{\Pi^{-1}(1)}, \mathbf{X}^M_{\Pi^{-1}(2)}, \cdots, \mathbf{X}^M_{\Pi^{-1}(N)}) \\ &= (\mathbf{Y}_1^M, \mathbf{Y}_2^M, \cdots, \mathbf{Y}_N^M),\end{aligned}$$

where Perm(.) shows the applied permutation function. Then,

$$\mathbf{Y}^M_{\Pi^{(N)}(i)} = \mathbf{X}_i^M = (X_i(1), X_i(2), \cdots, X_i(M))^T .$$

B. Markov Chain Model

Assume there is an area with r possible locations that users can go to. We use a Markov chain MC with r states to model movements of each user, Figure 1. We define E, the set of edges in the this Markov chain such that $(i, j) \in E$ if there exists an edge from i to j with probability $p_{ij} > 0$.

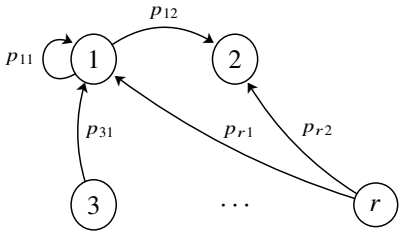

Fig. 1. Markov chain model MC with r states and $|E|$ number of edges.

We assume that this Markov chain gives the movement pattern of each user and what differentiates between users is their transition probabilities. That is, for fixed i and j, two different users could have two different transition probabilities, p_{ij}.

Note that when we are determining p_{ij}'s, there are d degrees of freedom which $d = |E| - r$. This is because for all i's, we must have

$$\sum_{j=1}^{r} p_{ij} = 1.$$

Thus, the Markov chain of the user u is completely determined by d values from p_{ij}'s which we show them as

$$\mathbf{P}^u = (p_1^u, p_2^u, \cdots, p_d^u)$$

and $\mathbf{P}^u$ is known to the adversary for $u = 1, 2, \cdots, N$. We define E_d the set of d edges whose p_{ij}'s belong to $\mathbf{P}^u$.
Let $R_p \subset \mathbb{R}^d$ be the range of acceptable values for $(p_1, p_2, \cdots, p_d)$, for example in Figure 2, for $i = 1, 2, 3$, we have

$$R_p = \{(p_1, p_2, p_3) \in \mathbb{R}^3 : 0 \le p_i \le 1, p_1 + p_2 \le 1\}.$$

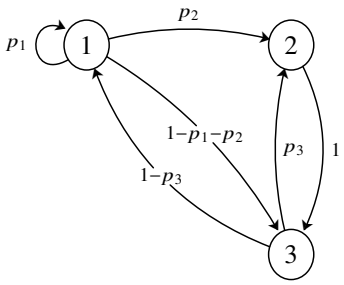

Fig. 2. Three states Markov chain example

The statistical properties of each user is completely determined since the adversary knows the Markov chain of each user. The adversary wants to be able to distinguish between users by having M observations per user and also knowing $\mathbf{P}^u$'s.

In this model we have N users and we assume that $(p_1, p_2, \cdots, p_d)$ are drawn independently from a d-dimensional continuous density function, $f(p)$, such that $\exists\ \delta_1, \delta_2 > 0$, $\delta_1 < f(p) < \delta_2$ for $p \in R_p$ and $f(p) = 0$ otherwise.

We now state and prove the main theorem that gives the condition for perfect location privacy for a user in the above setting.

Theorem 1. For an irreducible, aperiodic Markov chain with r states and $|E|$ edges, if $M = cN^{\frac{2}{|E|-r}-\alpha}$, where $c > 0$ and $\alpha > 0$ are constants, then

$$\lim_{N\to\infty} I(X_1(k); \mathbf{Y}^M) = 0, \quad \forall k \in \mathbb{N}, \tag{1}$$

i.e., user 1 has perfect location privacy.

In the rest of the paper, we use notation $M_i^{(k)}$ as user k's number of visits to state i and also $M_{ij}^{(k)}$ as number of transitions that user k has made from state i to j.

To prove theorem 1 we first introduce two lemmas.

Lemma 1. Given M_{ij}'s for all users which $(i, j) \in E$, then $\Pi^{(N)}$ and $\mathbf{Y}^M$ are independent.

Proof. (Sketch) This is the immediate result of the fact that M_{ij}'s are sufficient statistics for transition probabilities of the Markov chain. Using this, we conclude given that we know M_{ij} for all $(i, j) \in E$, no additional information about Π is learned by observing $\mathbf{Y}^M$. Specifically, if we let

$$\mathbf{M} = \left[M_{ij}\right]_{(i,j)\in E},$$

then

$$f(\mathbf{Y}^M|\mathbf{M}, \Pi) = f(\mathbf{Y}^M|\mathbf{M}).$$

□

Lemma 2. Consider user u with the given transition probabilities $\mathbf{P}^u$. Let $M_i^{(u)}$ be the number of visits to state i and π_i^u be the limiting probability of being at state i. Assume that $(i, j), (k, l) \in E$, then as $M \to \infty$,

(a) $\frac{M_i^{(u)}}{M} \to \pi_i^u$

(b) $\frac{M_{ij}^{(u)} - p_{ij}^u M_i^{(u)}}{\sqrt{M_i^{(u)} p_{ij}^u (1-p_{ij}^u)}} \xrightarrow{d} N(0, 1)$

(c) $\lim_{M\to\infty} \left|\mathrm{Cov}(\frac{M_{ij}^{(u)}}{\sqrt{M}}, \frac{M_{kl}^{(u)}}{\sqrt{M}})\right|$ exists and is finite

(d) $\frac{M_{ij}^{(u)} - p_{ij} M_i^{(u)}}{\sqrt{M_i^{(u)} p_{ij}(1-p_{ij})}}$, $\frac{M_{kl}^{(u)} - p_{kl} M_k^{(u)}}{\sqrt{M_k^{(u)} p_{kl}(1-p_{kl})}}$ are jointly Gaussian

where $\xrightarrow{d}$ shows the convergence in distribution.

Proof. (Sketch) Part (a) is the result of Markov chain convergence theorem which shows that any finite, irreducible, aperiodic Markov chain has a unique stationary distribution which is equal to the limiting distribution. Part (b) is the result of the Central Limit Theorem (CLT). To show parts (c) and (d), note that for a fixed i and given $M_i^{(u)}$, the random variable $M_{ij}^{(u)}$'s have multinomial distribution. Thus, as $M_i^{(u)} \to \infty$, they converge to a jointly Gaussian distribution. □

Proof of Theorem 1.

To prove theorem 1, let $\mathbf{P}^1 = (p_1^1, p_2^1, \cdots, p_d^1)$. Define the set $B^{(N)}$ such that

$$B^{(N)} = \{(x_1, \cdots, x_d) \in \mathbb{R}^d : p_i^1 - \epsilon \le x_i \le p_i^1 + \epsilon, i = 1, \cdots, d\},$$

where

$$\epsilon = \frac{1}{N^{\frac{1}{d}-\theta}}$$

and $0 < \theta < \frac{\alpha}{2}$ which α was given in $M = cN^{\frac{2}{|E|-r}-\alpha}$.

Let $J^{(N)} = \{u : (p_1^u, p_2^u, \cdots, p_d^u) \in B^{(N)}\}$ and note that $\forall u \in \{1, \cdots, N\}$,

$$\delta_1(2\epsilon)^d \le p(u \in J^{(N)}) \le \delta_2(2\epsilon)^d.$$

Thus, there exists $\delta > 0$, such that $\delta_1 < \delta < \delta_2$ and we can write

$$P(u \in J^{(N)}) = \delta(2\epsilon)^d, \text{for all u.}$$

Thus $|J(N)| \sim Bin(N, \delta(2\epsilon)^d)$. For expected value and variance of $|J^{(N)}|$ we have

$$E[|J^{(N)}|] = \delta N(2\epsilon)^d = \delta 2^d N^{d\theta} \triangleq \mu,$$

$$Var(|J^{(N)}|) = \delta N(2\epsilon)^d(1 - \delta(2\epsilon)^d) = \delta 2^d N^{d\theta}(1 - \delta 2^d N^{d\theta}),$$

and as $N \to \infty$, $Var(|J^{(N)}|) \approx \mu$.

Using Chebyshev's inequality

$$P\left\{\left||J^{(N)}| - E[|J^{(N)}|]\right| > \frac{\mu}{2}\right\} < \frac{Var(|J^{(N)}|)}{\frac{\mu^2}{4}}$$

$$P\left\{\left||J^{(N)}| - E[|J^{(N)}|]\right| > \frac{\mu}{2}\right\} < \frac{\mu}{\frac{\mu^2}{4}} = \frac{4}{\mu} \to 0,\ as\ N \to \infty.$$

In particular with high probability,

$$|J^{(N)}| \to \infty,\ as\ N \to \infty.$$

Now note that, given that $u \in J^{(N)}$, as $N \to \infty$ then $\epsilon \to 0$ and we can write

$$(p_1^u, p_2^u, \cdots, p_d^u) \to (p_1^1, p_2^1, \cdots, p_d^1), \quad as\ N \to \infty.$$

This suggests that it is difficult for the adversary to distinguish between these users. Now, to finish the proof of Theorem 1, we show that for the adversary, all users in the $J^{(N)}$ have the same asymptotic Markov chain distribution. More specifically, we have the following lemma.

Lemma 3. For all $u \in J^{(N)}$, and $(i, j) \in E_d$, we have

(a) $\frac{M_{ij}^{(u)} - p_{ij}^1 M_i^{(u)}}{\sqrt{M_i^{(u)} p_{ij}^1 (1-p_{ij}^1)}} \xrightarrow{d} N(0,1)$.

(b) $\lim_{M\to\infty} \left|\mathrm{Cov}(\frac{M_{ij}^{(u)}}{\sqrt{M}}, \frac{M_{kl}^{(u)}}{\sqrt{M}})\right|$ are the same for all users in $J^{(N)}$.

Proof. To prove part (a), note that by definition of set $J^{(N)}$, for all $u \in J^{(N)}$

$$\left|p_{ij}^{(u)} - p_{ij}^1\right| < \epsilon = \frac{1}{N^{\frac{1}{d}-\theta}},$$

As $N \to \infty$, $p_{ij}^{(u)} \to p_{ij}^1$ and also $\pi_i^{(u)} \to \pi_i^1$. Using lemma 2 and this we can write

$$\frac{M_i^{(u)}}{M} \to \pi_i^{(1)}.$$

Now we define W_n as

$$W_n = \frac{M_{ij}^{(u)} - p_{ij}^1 M_i^{(u)}}{\sqrt{M_i^{(u)} p_{ij}^1 (1-p_{ij}^1)}}$$

and we can write that as

$$W_n = \frac{M_{ij}^{(u)} - p_{ij}^u M_i^{(u)} + p_{ij}^u M_i^{(u)} - p_{ij}^1 M_i^{(u)}}{\sqrt{M_i^{(u)} p_{ij}^1 (1-p_{ij}^1)}} =$$

$$\frac{M_{ij}^{(u)} - p_{ij}^u M_i^{(u)}}{\sqrt{M_i^{(u)} p_{ij}^u (1-p_{ij}^u)}} \cdot \frac{\sqrt{p_{ij}^u (1-p_{ij}^u)}}{\sqrt{p_{ij}^1 (1-p_{ij}^1)}} + \frac{M_i^{(u)}}{\sqrt{M_i^{(u)}}} \cdot \frac{p_{ij}^{(u)} - p_{ij}^1}{\sqrt{p_{ij}^1 (1-p_{ij}^1)}}.$$

Since $p_{ij}^{(u)} \to p_{ij}^1$, $\frac{p_{ij}^{(u)} - p_{ij}^1}{\sqrt{p_{ij}^1 (1-p_{ij}^1)}} \to 1$ and using Lemma 2 we have that

$$\frac{M_{ij}^{(u)} - p_{ij}^u M_i^{(u)}}{\sqrt{M_i^{(u)} p_{ij}^u (1-p_{ij}^u)}} \xrightarrow{d} N(0,1).$$

Now we need to show that as $N \to \infty$,

$$\frac{M_i^{(u)}}{\sqrt{M_i^{(u)}}} \cdot \frac{p_{ij}^{(u)} - p_{ij}^1}{\sqrt{p_{ij}^1 (1-p_{ij}^1)}} \to 0.$$

We define B as,

$$B = \frac{\sqrt{M_i^{(u)}}}{\sqrt{p_{ij}^1 (1-p_{ij}^1)}} (p_{ij}^{(u)} - p_{ij}^1).$$

Since $\frac{M_i^{(u)}}{M} \to \pi_i^{(1)}$ we can write,

$$|B| < \frac{1}{\sqrt{p_{ij}^1 (1-p_{ij}^1)}} \epsilon \sqrt{\pi_i^{(1)} M} =$$

$$\lambda \sqrt{c.N^{\frac{2}{d}-\alpha}} . \frac{1}{N^{\frac{1}{d}\theta}} = \gamma . N^{\theta - \frac{\alpha}{2}}$$

where λ, γ are constants. Note $0 < \theta < \frac{\alpha}{2}$, so as $N \to \infty$,

$$\gamma . N^{\theta - \frac{\alpha}{2}} \to 0,$$

which means that,

$$\frac{M_i^{(u)}}{\sqrt{M_i^{(u)}}} \cdot \frac{p_{ij}^{(u)} - p_{ij}^1}{\sqrt{p_{ij}^1 (1-p_{ij}^1)}} \to 0.$$

and we can say that $W_n \xrightarrow{d} N(0,1)$. Part (b) can be shown similarly by applying the CLT. □

At this point, we have all the components needed to prove Theorem 1. We have basically established that there exists a large number ($|J^{(N)}|$) of users who are indistinguishable from user 1. Now, by using exactly the same reasoning as Lemma 4 and Lemma 5 in [1] and combining them with lemmas 1,2 and 3, we can conclude that

$$\lim_{N\to\infty} I(X_1(k); \mathbf{Y}^M) = 0, \quad \forall k \in \mathbb{N},$$

i.e., user 1 has perfect location privacy.

IV. Simulation

Here, we provide some simulation results that verify the result in Theorem 1. We consider a network with N users and r locations. Possible path of each user can be modeled as an irreducible, aperiodic Markov chain with r states and $|E|$ number of edges. After obtaining M observations per user, the adversary estimates transition probabilities $p_{ij}^{(u)}$ as $\frac{M_{ij}^{(u)}}{M_i}$ and by using nearest neighbor decoding in $\mathbb{R}^d$, she matches a user to the observed paths.

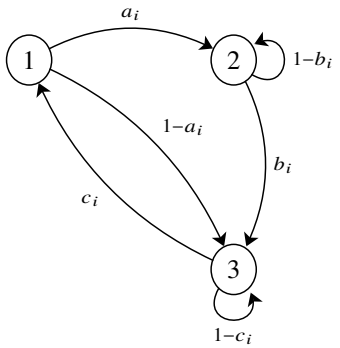

Fig. 3. The Markov chain MC which models of users' path.

We see that if the adversary's number of observations, M, is more than $O(N^{\frac{2}{|E|-r}})$, then the adversary's error probability

(when adversary fails to map user with pseudonym $\Pi(i)$ to user i) goes to zero. On the other hand, if the number of observations is much smaller, then the error probability goes to one suggesting that users might have perfect location privacy.

In our simulations we consider $r = 3$ and $M = \beta N^{\frac{2}{|E|-r}}$. We model each user's path as a Markov chain MC shown in figure 3. Since in this model $|E| = 6$ we can write $M = \beta N^{\frac{2}{3}}$.

In order to have N unique users, we generate each user's transition probabilities $\mathbf{P}^u$ at random based on a uniform distribution on R_p and we consider them known to the adversary. For a fixed $\beta = 5$, figure 4 shows that as N increases, the error probability of the adversary converges to a fix positive value. We have repeated this for different values of β and have observed the same effect. This is consistent with our result that $M = O(N^{\frac{2}{|E|-r}})$ is the threshold for perfect privacy.

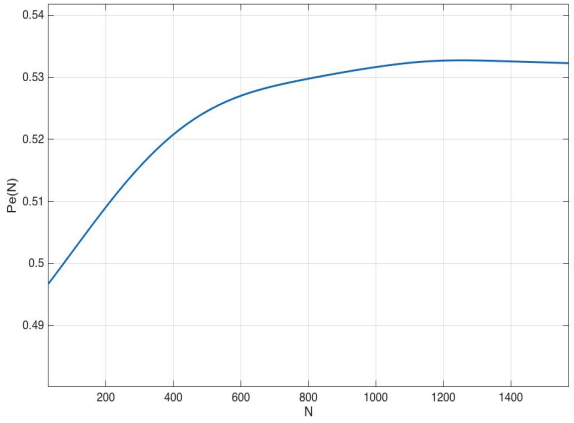

Fig. 4. $P_e(N)$ vs. N for Markov chain MC with $\beta = 5$.

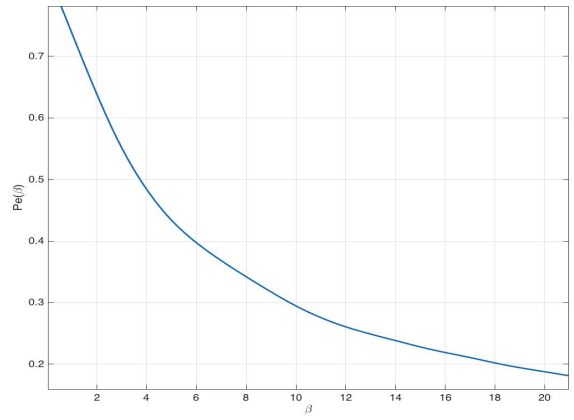

Fig. 5. $P_e(\beta)$ vs. β for Markov chain MC with $N = 500$.

Next, we fix $N = 500$. Simulation results in figure 5 shows that as β grows, the adversary's error probability goes to zero which shows that the adversary maps users with less error probability. On the other hand, as β becomes smaller, the error probability approaches 1. These results are consistent with the our main result that users have perfect privacy if the adversary obtains less than $O(N^{\frac{2}{|E|-r}})$ observations per user.

V. Conclusion

We provided an information theoretic definition for perfect location privacy. Using Markov chains, we modeled each user's movements over time. We proved that for such a user, perfect location privacy is achievable if pseudonym of the user is changed before $O(N^{\frac{2}{|E|-r}})$ observations is made by the adversary. Simulated results for different scenarios are also consistent with this result.

References

[1] Z. Montazeri, A. Houmansadr, and H. Pishro-Nik, "Defining perfect location privacy using anonymization," in *2016 Annual Conference on Information Science and Systems (CISS) (CISS 2016)*, Princeton, USA, Mar. 2016.

[2] R. Shokri, G. Theodorakopoulos, C. Troncoso, J.-P. Hubaux, and J.-Y. Le Boudec, "Protecting location privacy: optimal strategy against localization attacks," in *Proceedings of the 2012 ACM conference on Computer and communications security*. ACM, 2012, pp. 617–627.

[3] M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking," in *Proceedings of the 1st international conference on Mobile systems, applications and services*. ACM, 2003, pp. 31–42.

[4] B. Hoh, M. Gruteser, H. Xiong, and A. Alrabady, "Preserving privacy in gps traces via uncertainty-aware path cloaking," in *Proceedings of the 14th ACM conference on Computer and communications security*. ACM, 2007, pp. 161–171.

[5] M. Wernke, P. Skvortsov, F. Dürr, and K. Rothermel, "A classification of location privacy attacks and approaches," *Personal and Ubiquitous Computing*, vol. 18, no. 1, pp. 163–175, 2014.

[6] Y. Cai and G. Xu, "Cloaking with footprints to provide location privacy protection in location-based services," Jan. 1 2015, uS Patent App. 14/472,462. [Online]. Available: https://www.google.com/patents/US20150007341

[7] H. Kido, Y. Yanagisawa, and T. Satoh, "An anonymous communication technique using dummies for location-based services," in *Pervasive Services, 2005. ICPS'05. Proceedings. International Conference on*. IEEE, 2005, pp. 88–97.

[8] H. Lu, C. S. Jensen, and M. L. Yiu, "Pad: privacy-area aware, dummy-based location privacy in mobile services," in *Proceedings of the Seventh ACM International Workshop on Data Engineering for Wireless and Mobile Access*. ACM, 2008, pp. 16–23.

[9] R. Shokri, G. Theodorakopoulos, J.-Y. Le Boudec, and J.-P. Hubaux, "Quantifying location privacy," in *Security and Privacy (SP), 2011 IEEE Symposium on*. IEEE, 2011, pp. 247–262.

[10] R. Shokri, G. Theodorakopoulos, G. Danezis, J.-P. Hubaux, and J.-Y. Le Boudec, "Quantifying location privacy: the case of sporadic location exposure," in *Privacy Enhancing Technologies*. Springer, 2011, pp. 57–76.

[11] L. Sweeney, "k-anonymity: A model for protecting privacy," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 05, pp. 557–570, 2002.

[12] A. R. Beresford and F. Stajano, "Location privacy in pervasive computing," *IEEE Pervasive computing*, no. 1, pp. 46–55, 2003.

[13] J. Freudiger, R. Shokri, and J.-P. Hubaux, "On the optimal placement of mix zones," in *Privacy enhancing technologies*. Springer, 2009, pp. 216–234.

[14] Z. Ma, F. Kargl, and M. Weber, "A location privacy metric for v2x communication systems," in *Sarnoff Symposium, 2009. SARNOFF'09. IEEE*. IEEE, 2009, pp. 1–6.

[15] T. Li and N. Li, "On the tradeoff between privacy and utility in data publishing," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2009, pp. 517–526.

[16] S. Salamatian, A. Zhang, F. du Pin Calmon, S. Bhamidipati, N. Fawaz, B. Kveton, P. Oliveira, and N. Taft, "How to hide the elephant-or the donkey-in the room: Practical privacy against statistical inference for large data." in *GlobalSIP*, 2013, pp. 269–272.

[17] F. P. Calmon, A. Makhdoumi, and M. Médard, "Fundamental limits of perfect privacy," in *Information Theory (ISIT), 2015 IEEE International Symposium on*. IEEE, 2015, pp. 1796–1800.