IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 6, 2018, accepted June 5, 2018, date of publication June 18, 2018, date of current version July 25, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2847720

An Optimal Pufferfish Privacy Mechanism for Temporally Correlated Trajectories

LU OU[1], (Student Member, IEEE), ZHENG QIN[1], (Member, IEEE), SHAOLIN LIAO[2,3], (Senior Member, IEEE), HUI YIN[1,4], AND XIAOHUA JIA[5], (Fellow, IEEE)

[1]College of Computer Science and Electronic Engineering, Hunan University, Changsha 410082, China
[2]Argonne National Laboratory, Lemont, IL 60439, USA
[3]Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, IL 60616, USA
[4]College of Computer Engineering and Applied Mathematics, Changsha University, Changsha 410022, China
[5]Department of Computer Science, City University of Hong Kong, Hong Kong

Corresponding author: Zheng Qin (zqin@hnu.edu.cn)

This work was supported in part by the Key Project of National Science Foundation of China under Grant 61732022, in part by the National Science Foundation of China under Grant 61472131, Grant 61472132, and Grant 61772191, in part by the Science and Technology Key Projects of Hunan Province under Grant 2015TP1004, Grant 2015SK2087, Grant 2015JC1001, and Grant 2016JC2012, in part by the Natural Science Foundation of Hunan Province under Grant 2017JJ2292, in part by the Outstanding Youth Research Project of the Provincial Education Department of Hunan under Grant 17B030, and in part by the Science and Technology Planning Project of Changsha under Grant K1705018.

ABSTRACT Temporally correlated trajectories are ubiquitous, and it has been a challenging problem to protect the temporal correlation from being used against users' privacy. In this paper, we propose an optimal Pufferfish privacy mechanism to achieve better data utility while providing guaranteed privacy of temporally correlated daily trajectories. First, a Laplace noise mechanism is realized through geometric sum of noisy Fourier coefficients of temporally correlated daily trajectories. Then, we prove that the proposed noisy Fourier coefficients' geometric sum satisfies Pufferfish privacy, *i.e.*, the so-called FGS-Pufferfish privacy mechanism. Furthermore, we achieve better data utility for a given privacy budget by solving a constrained optimization problem of the noisy Fourier coefficients via the Lagrange multiplier method. What is more, a rigorous mathematical formula has been obtained for the Fourier coefficients' Laplace noise scale parameters. At last, we evaluate our FGS-Pufferfish privacy mechanism on both simulated and real-life data and find that our proposed mechanism achieves better data utility and privacy compared with the other state-of-the-art existing approach.

INDEX TERMS Fourier coefficients, geometric sum, Lagrange multiplier method, Pufferfish privacy, temporally correlated trajectories.

I. INTRODUCTION

Currently, many real-world applications, such as maps, points of interest searching and taxi reservation, generate, store and process a large amount of temporally correlated trajectories. Releasing these trajectories makes it possible for researchers to predict users' behavior trend [1] and monitor traffic [2] by utilizing temporal correlations of users, *i.e.*, relations between locations at different timestamps.

Although temporal correlations of users' trajectories are useful to the researchers and many service applications such as travel recommendations [1], they may cause a privacy risk [3]. With background knowledge, adversaries may mine individuals' privacy through the temporal correlations of their trajectories. For example, as shown in Figure 1, once trajectories are released to untrusted service providers, they may compute the temporal correlation of a user's trajectories. Then the adversaries could infer the user's mobility pattern, *e.g.*, Alice, with their background knowledge about Alice's temporal correlation. Furthermore, they may analyze her privacy such as home address and work address.

The problem of privacy-preserving trajectories releasing has attracted extensive interests of many researchers. As direct solutions, dummy trajectories [4], suppression [5], [6] and k-anonymity [7] have been adopted to

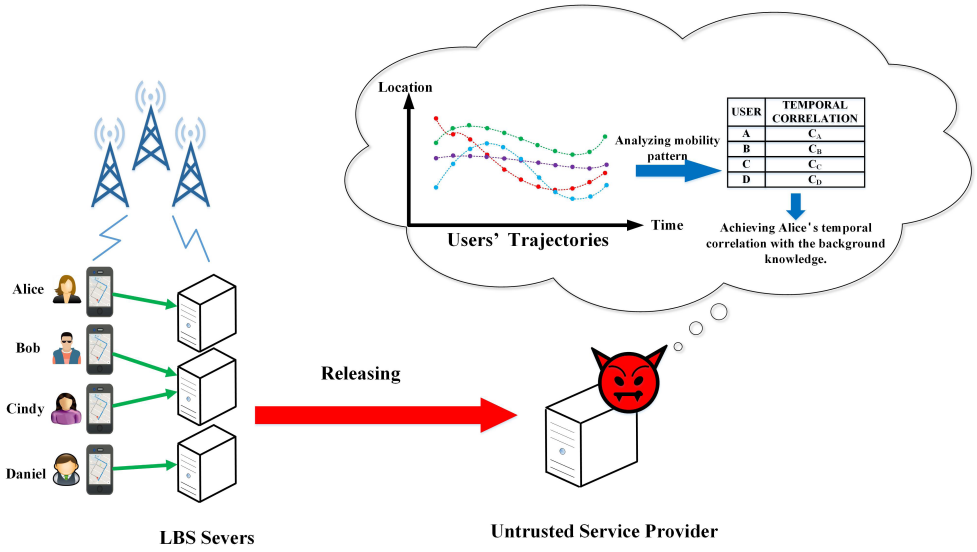


FIGURE 1. Temporal correlation inference attacking.

protect trajectory privacy. Unfortunately, they cannot resist composition attack [8], deFinetti attack and background knowledge attack [9]. To overcome this, differential privacy [10] is considered to achieve privacy-preserving trajectory publication [11], [12] because adversaries cannot judge whether individuals are in the database or not. Nevertheless, these solutions do not consider the issue of temporal correlations of trajectories.

In order to protect individuals' privacy under the assumption that there are temporal correlations within trajectories, several works have attempted to consider the relations between locations for differentially private trajectory releasing. For the privacy-preserving aggregate information releasing, first, the prefix tree [13] is adopted to achieve such goal. Then, the n-gram model [14] is introduced to protect correlation privacy. Recently, Markov model [15] is utilized to achieve differentially private correlation protection for trajectories. For a single trajectory releasing, the sum of many copies of Gaussian white noise [3] is adopted to generate correlated Laplace noise for protecting trajectories' temporal correlations, *i.e.*, the so-called Correlated Time-Series (CTS) privacy protection. Furthermore, based on Markov model, Pufferfish privacy [16] is adopted to protect the temporal correlations within trajectories [17]. Although these solutions can protect temporal correlations efficiently, they do not consider to optimize data utility for a given privacy budget.

It has been a task of great challenge to handle such individuals' privacy induced by the temporal correlations. On one hand, Differential Privacy (DP) cannot be directly applied to temporally correlated trajectories. Although a Pufferfish privacy mechanism can handle such kind of trajectories, to the best of our knowledge, there is no rigorous mathematics formulation of a Pufferfish privacy mechanism dedicated to the temporal correlation privacy problem. On the other hand, to have an efficient Pufferfish privacy mechanism, it is preferred that a rigorous mathematics formula of the noise scale parameters can also be obtained to achieve the optimal data utility with a guaranteed privacy. To fill in such gap, in this paper, we propose a Laplace noisy mechanism, satisfying Pufferfish privacy, for protecting the individuals' privacy against adversarial inferencing through their trajectories' temporal correlations.

The main contributions of this paper are summarized below.

- We propose a noise mechanism for efficient generation of Laplace noise via the Fourier coefficients' geometric sum to achieve ε-Pufferfish privacy, *i.e.*, the FGS-Pufferfish privacy mechanism.
- We present an analytical formula of the optimized the Fourier coefficients noise for the constrained optimization problem of achieving a better data utility for a given privacy budget. Then, we provide theoretical analysis of the data utility and privacy, as well as the posterior-to-prior knowledge gain of an adversary.
- We evaluate the proposed FGS-Pufferfish privacy mechanism over both simulated and real-life datasets. The experimental results demonstrate that our proposed mechanism does achieve better privacy and data utility than the state-of-the-art existing approach.

The remainder of this paper is organized as follows. In Section II, we review the related work in the literature. In Section III, we formally define our problem, and then introduce some background knowledge, including Pufferfish privacy, discrete Fourier transform and its inverse transform, geometric sum, and the constrained optimization problem. Then, in Section IV, we explain the relation between the mobility pattern and temporal correlations,

and mathematically quantify the temporal correlation of a trajectory and the corresponding temporal correlation's discrete Fourier transform. In Section V, we introduce our proposed FGS-Pufferfish privacy mechanism. Furthermore, data utility, privacy analyses and the adversary's posterior-to-prior knowledge gain are depicted in Section VI. The privacy-preserving trajectory releasing algorithm based on the FGS-Pufferfish privacy mechanism is outlined in Section VII. In Section VIII, the experimental performance evaluations are shown. Finally, Section IX concludes this paper.

II. RELATED WORK

Releasing temporally correlated raw trajectories poses a serious privacy problem, due to potential sensitive information leak linked to the temporal correlation. In this section, we review some literature about the privacy-preserving trajectories releasing.

Gao *et al*. [18], [19] reviewed different methods for trajectory privacy. At the beginning, researchers proposed dummy trajectories [4], suppression technique [5], [6], techniques based on k-anonymity [7] and their variants to protect privacy in trajectories. Unfortunately, these approaches cannot prevent inferencing attack well if the locations are in sensitive areas. Researchers also adopted differential privacy to protect trajectory privacy. Rastogi and Nath [11] proposed the Fourier Perturbation Algorithm (FPA) using Fourier transform, without considering the temporal correlations in a trajectory. Jiang *et al*. [12] proposed a noisy position mechanism for each position, satisfying (ε, δ)-differential privacy, and a noisy coordinate mechanism whose Laplace scale of the longitude is the same as that of the latitude, satisfying ε-differential privacy, respectively. Then they proposed an ε-differentially private exponential mechanism through sampling suitable distance and director for trajectory publication. Riboni and Bettini [20] proposed a trajectory privacy protection in a context-aware recommending system by combining the (L, j)-density with ε-differential privacy. Quan *et al*. [21] proposed a trajectory obfuscation mechanism based on the Laplace mechanism. In this mechanism, a polar Laplace noise is added on the trajectories. Zhang *et al*. [22] proposed a noise generation strategy based on the time-series pattern in order to protect individual privacy within the cloud framework. Cao and Yoshikawa [23] proposed an ℓ-trajectory privacy model to protect a trajectory whose length is ℓ. This model satisfies ε-differential privacy for releasing real-time statistics of trajectory streams. Under the assumption of no temporal correlation, Hua *et al*. [9] designed an exponential mechanism to select a group divided by the distances between locations at each timestamp and proposed a Laplace mechanism of noisy counts for differentially private trajectory publication. Furthermore, Li *et al*. [24] proposed a differentially private location generation algorithm and a bounded Laplace mechanism for trajectories releasing. Besides, Bindschaedler and Shokri [25] generated fake trajectories by utilizing the correlations between locations for privacy-preserving trajectory releasing. Moreover, these approaches also do not consider trajectories' temporal correlations which could leak individual privacy.

However, an adversary may utilize the temporal correlation to build a user's mobility pattern and thus infer the user's trajectory [1], [2]. There are several works dealing with such spatio-temporal correlation. On one hand, researches considered privacy-preserving aggregate information releasing. Chen *et al*. [13] proposed a data-dependent solution by recursively constructing a noisy prefix tree based on the existed trajectory data, for trajectory statistics publishing. However, with the growth of the prefix tree, this solution will lead to poor data utility. To solve this problem, Chen *et al*. [14] proposed a differentially private trajectory statistics publishing method by using the variable n-gram model. He *et al*. [26] proposed a trajectory systhesis method, called ''DPT'', according to the correlations between locations within a single trajectory. They considered individual movements' speeds and constructs prefix tree counts ensuring ε-differential privacy. Then, based on this approach, they presented a tool, called ''VisDPT'' [27], helping data curators understand privacy problems for data releasing and their proposed privacy protecting mechanism. Fan *et al*. [28] utilized road networking, overall population density and so on to model the correlations at per timestamps and adopted Quadtree to deal with the sparsity of trajectories for differentially private aggregate releasing, respectively. Furthermore, a set of novel techniques are based on the Markov assumption. Wang and Sinnott [15] proposed a private reference system by cluster-based anchor points under the X-order Markov assumption. The raw trajectories are discrete in this system. Then noisy calibrated trajectories are released by using differentially private prefix trees.

On the other hand, researchers generated privacy-preserving trajectories to protect temporal correlations. Wang and Xu [3] adopted Gaussian white noise to protect a trajectory's temporal correlation. Ou *et al*. [29] proposed differentially private trajectory releasing by using Hidden Markov model. To deal with the weakness of differential privacy, Song *et al*. [17] proposed the Markov Quilt Mechanism (MQM) under the Pufferfish privacy framework to protect the spatio-temporal correlation within a single trajectory. Although these works consider the spatio-temporal correlation within a single-user trajectory, the data utility optimization is not achieved for a given privacy budget.

In sum, although there are several literatures dealing with the privacy problem of temporal correlations, none of them presents a rigorous mathematical Laplace noise mechanism under the Pufferfish privacy framework that is proved to be ε-differentially private. What's more, to the best of our knowledge, there is no rigorous analytical formula of the noise scale parameters to achieve the optimal data utility, for a given privacy budget. Thus in this paper, we propose such Laplace noise mechanism through adding noise in the daily trajectory's Fourier coefficients, satisfying ε-Pufferfish privacy, in order to protect temporal correlations against inferencing attacks.

III. PRELIMINARIES

In this section, we first present the statement of our problem and the basic idea of our mechanism. Then, we introduce some basic concepts, including Pufferfish privacy, Discrete Fourier Transform (DFT) and Inverse Discrete Fourier Transform (IDFT), geometric sum, and the constrained optimization problem. To start, **Table 1** lists some key variables used across this paper with their explanations.

TABLE 1. Notations and definitions.

Symbol	Description
t_n	The n-th time slot of a day.
(X_n, Y_n)	A daily location at t_n of a user.
$\mathbb{T}$	A daily trajectory: $\{(X_n, Y_n)\,\|n = 0, 1, \cdots, N-1\}$.
(x_n, y_n)	A modified location: $\left(X_n - \overline{X}_n, Y_n - \overline{Y}_n\right)$.
$\mathcal{F}_k$	The k-th Fourier coefficient of a user's modified daily. trajectory $\{(x_n, y_n)\|n = 0, 1, \cdots, N-1\}$.
$\mathbb{C}$	A temporal correlation of a user's modified daily trajectory $\{(x_n, y_n)\|n = 0, 1, \cdots, N-1\}$.
$\mathcal{S}_k$	The k-th Fourier coefficient of a user's temporal correlation $\mathbb{C}$.
$\mathbf{E}\{X\}$	Obtain mean of a random variable X.
$\mathbf{Var}\{X\}$	Obtain variance of a random variable X.
$\overline{X}$	The mean value of a random variable X.
σ_X	The standard deviation of a random variable X.
b_X	A Laplace scale parameter of a random variable X.
$\mathcal{U}_L$	A location utility of the noisy modified daily trajectories.
$\mathcal{U}_\mathbb{C}$	A correlation utility of the noisy modified daily trajectories.
w	The weight of the location utility $\mathcal{U}_L$.
$\mathcal{U}$	A weighted data utility: $w\mathcal{U}_L + (1-w)\mathcal{U}_\mathbb{C}$.
ε	Privacy budget of all users' temporal correlations set.

A. PROBLEM STATEMENT AND BASIC IDEA

Our goal is to protect the temporal correlation of a user's daily trajectories from being related to a user's daily trajectories pattern and thus the user's privacy, when combined with other prior knowledge about the user. First, let's define the daily trajectory of a user as follows.

Definition 1 (Daily Trajectory): The daily trajectory of a user u_m on the d-th day, denoted as $\mathbb{T}$, is a sequence of locations at N times slots,

$$\mathbb{T} = \left\{(X_n, Y_n)\,\middle|n = 0, 1, \ldots, N-1\right\},$$

where X_n and Y_n are the longitude and latitude of a user u_m at the n-th time slot on the d-th day, respectively; and $m \in \{1, 2, \ldots, M\}$ and $d \in \{1, 2, \ldots, D\}$.

A daily trajectory database, denoted by $\mathcal{T}$, contains M users' daily trajectories over D days. When adversaries obtain the database $\mathcal{T}$, they may compute the temporal correlation of every user through ensemble average of the user's daily trajectories over all D days; then the adversaries may compare it with the prior knowledge of a user's temporal correlation. Based on the similarity and other background knowledge, they could identify which temporal correlation belongs to a certain user. Therefore the temporal correlation of a user's daily trajectories should be protected in order to avoid such inferencing attack.

In this paper, we first quantify the temporal correlation in a rigorous mathematics way. Then we propose the FGS-Pufferfish privacy mechanism by adding noise to the Fourier coefficients through geometric sum. At last, we obtained the rigorous formula of the optimal noisy Fourier coefficients by solving the constrained optimization problem, *i.e.*, achieving the optimal data utility for a given privacy budget.

For the rest of this paper, because the mean or the ensemble average of a user's daily trajectories does not affect our problem treatment, we will only consider the following modified daily trajectory with its mean subtracted: $\{(x_n, y_n)\,|n = 0, 1, \ldots, N-1\}$, where $(x_n, y_n) = \left(X_n - \bar{X}_n, Y_n - \bar{Y}_n\right)$, where $(\bar{X}_n, \bar{Y}_n)$ is the ensemble average of the n-th time-slot locations over all D days, *i.e.*, $(\bar{X}_n, \bar{Y}_n) = \mathbf{E}_d\left\{\left(X_n^{(d)}, Y_n^{(d)}\right)\right\}$ $(d \in \{1, 2, \ldots, D\})$. Also, we will assume that $\{x_n|n = 0, 1, \ldots, N-1\}$ and $\{y_n|n = 0, 1, \ldots, N-1\}$ are statistically independent and can be treated individually. What's more, for simplicity, we take the modified longitude daily trajectory $\{x_n|n = 0, 1, \ldots, N-1\}$ as an example to explain our proposed mechanism.

B. BASIC CONCEPTS

Before we go into the details of the temporal correlation privacy issue, let's introduce some basic concepts first, which includes Pufferfish privacy, DFT/IDFT, geometric sum, and the constrained optimization problem.

1) PUFFERFISH PRIVACY

Pufferfish privacy [16] is a mathematics mechanism for privacy and its definition is given as below,

Definition 2 (Pufferfish Privacy): A random mechanism M is said to be ε-Pufferfish private in a framework $(\mathcal{C}, \mathcal{Q}, \Theta)$ if for the data $\mathbb{C}$ drawn from all possible belief distributions $\boldsymbol{\theta} \in \Theta$ of an adversary, the following condition is satisfied for all secret pairs $\left(\mathbb{C}^{(u)}, \mathbb{C}^{(v)}\right) \in \mathcal{Q}$, and all $\mathbb{C}' \in Range\{M(\mathbb{C} \in \mathcal{C})\}$,

$$\exp(-\varepsilon) \leq \frac{Pr\left(M(\mathbb{C}')\middle|\mathbb{C}^{(u)}, \boldsymbol{\theta}\right)=\mathbb{C}'}{Pr\left(M(\mathbb{C}')\middle|\mathbb{C}^{(v)}, \boldsymbol{\theta}\right)=\mathbb{C}'} \leq \exp(\varepsilon).$$

2) DISCRETE FOURIER TRANSFORM

Now let's look at the DFT of a user's daily trajectory.

Definition 3 (Discrete Fourier Transform): The discrete Fourier transform transforms a user's daily trajectory $\left\{x_n\middle|n = 0, 1, \ldots, N-1\right\}$ into a set of sine and cosine waves of different frequencies and corresponding Fourier coefficients $\left\{\mathcal{F}_k\middle|k = 0, 1, \ldots, K-1\right\}$ which is defined as follows,

$$\mathcal{F}_k = \sum_{n=0}^{N-1} x_n W_{n,k}^{-1}; \quad W_{n,k}^{-1} = \exp\left(-j\frac{2\pi}{N}nk\right),$$

where $j = \sqrt{-1}$; N and K are the total number of time slots and Fourier coefficients, respectively.

In this paper, we assume that both the real part $\mathcal{F}_k^r$ and the imaginary part $\mathcal{F}_k^i$ of the Fourier coefficient $\mathcal{F}_k$ follow the same Gaussian distribution with mean $\mu_k = 0$ and the

standard deviation of σ_{S_k}, and the probability distribution of $\{\mathcal{F}_k|k=0,1,\ldots,K-1\}$ is as follows,

$$Pr\left(\{\mathcal{F}_k|k=0,1,\ldots,K-1\}\right)=\prod_{k=0}^{K-1}\frac{\exp\left(-\frac{|\mathcal{F}_k|^2}{(\sigma_{S_k})^2}\right)}{2\pi(\sigma_{S_k})^2},$$

and its power spectrum $\mathcal{S}_k=|\mathcal{F}_k|^2$ follows the Chi-Square χ_2^2 distribution with a degree of 2, which is also the exponential distribution,

$$Pr\left(\{\mathcal{S}_k|k=0,1,\ldots,K-1\}\right)=\prod_{k=0}^{K-1}\frac{\exp\left(-\frac{S_k}{2\sigma_{S_k}^2}\right)}{2\sigma_{S_k}^2}. \quad (1)$$

On the contrary, a trajectory for a given Fourier coefficients set of length K can be expressed by the IDFT, and its definition is given as follows.

Definition 4 (Inverse Discrete Fourier Transform): The inverse discrete Fourier transform transforms the Fourier coefficients $\left\{\mathcal{F}_k|k=0,1,\ldots,K-1\right\}$ into the daily trajectory $\left\{x_n|n=0,1,\ldots,N-1\right\}$ which is defined as follows,

$$x_n=\sum_{k=0}^{K-1}\mathcal{F}_k W_{n,k};\quad W_{n,k}=\exp\left(j\frac{2\pi}{N}nk\right). \quad (2)$$

3) GEOMETRIC SUM

Geometric sum [30] can achieve noise Laplace distribution through sum of a random K-series of random variables $\left\{F_k|k=0,1,\ldots,K-1\right\}$ and its definition is as follows,

Definition 5 (Geometric Sum): Let $\left\{F_k|k=0,1,\cdots,K-1\right\}$ be a sequence of independent random variables (but not necessarily identically distributed), their geometric sum is defined as,

$$F=\sum_{k=0}^{K-1}F_k,$$

with K following the geometric distribution,

$$Pr(K)=(1-p)^{K-1}p=Geo(K;p).$$

Theorem 1: Suppose that $\mathbf{E}\{F_k\}=0$ and $\mathbf{Var}\{F_k\}=\sigma_{F_k}^2$, then the geometric sum F weakly follows the Laplace distribution [30],

$$Pr\left(F=\sum_{k=0}^{K-1}F_k\right)=Lap\left(F;0,\frac{\sigma_F}{\sqrt{2}}\right),$$

under the conditions of

$$\lim_{k\to\infty}k^{-\alpha}\sigma_{F_k}^2,\quad \text{for some}\quad 0<\alpha<1;$$

$$\sigma_F^2=\lim_{n\to\infty}\frac{1}{K}\sum_{k=0}^{K-1}\sigma_{F_k}^2>0 \text{ exists};$$

and for all $\xi>0$:

$$\lim_{p\to 0}\sum_{k=0}^{\infty}(1-p)^{k-1}p\,\mathbf{E}_{F_k}\left\{F_k^2;|F_k|\geq\xi p^{-1/2}\right\}=0.$$

4) THE CONSTRAINED OPTIMIZATION PROBLEM

The constrained optimization problem is a strategy of finding the local extrema (maxima and minima) of a function $f(b)$ subject to equality constraint $g(b)=0$. The constrained optimization problem is given below,

$$b^*=\underset{b}{\text{argmax}}\left\{f(b)|g(b)=0\right\},$$

with the following Lagrangian,

$$\begin{cases}\mathcal{L}(b|\lambda)=f(b)-\lambda g(b),\\ g(b)=0.\end{cases} \quad (3)$$

which can be solved by the Lagrange Multiplier (LM) method when the following conditions are satisfied,

$$\left.\frac{\partial\mathcal{L}}{\partial b}\right|_{(\lambda^*,b^*)}=\left.\frac{\partial\mathcal{L}}{\partial\lambda}\right|_{(\lambda^*,b^*)}=0.$$

IV. TEMPORAL CORRELATION AND MOBILITY PATTERN

In this section, we show that relation between the user's mobility pattern of a user and its temporal correlation $\mathbb{C}$ of the user's daily trajectories; then we show that its Fourier coefficients $\left\{\mathcal{S}_k|k=0,1,\ldots,K-1\right\}$ is closely related to the variance of the Fourier coefficients of a user over all D days, *i.e.*, $\left\{\mathbf{Var}_d\left\{\mathcal{F}_k^{(d)}\right\}|k=0,1,\ldots,K-1;d=1,2,\ldots,D\right\}$.

A. MOBILITY PATTERN

The user's mobility pattern can be described by the conditional probability of the next i-th location from the current n-th location, *i.e.*, $Pr\left(x_{n+i}|x_n\right)$. Also, for Markov process, the conditional probability of the current location depends only on its last location, *i.e.*, $Pr\big(x_{n+i}|x_{n+i-1},x_{n+i-2},\cdots,x_n\big)=Pr\left(x_{n+i}|x_{n+i-1}\right)$.

B. TEMPORAL CORRELATION

The temporal correlation $\mathbb{C}$ describes the relation among different locations within a user's daily trajectories $\{x_n|n=0,1,\ldots,N-1\}$ through the conditional transition probability $Pr\left(x_{n+i}|x_n\right)$, *i.e.*, the mobility pattern. It is the ensemble average of a user's daily trajectories $\{x_n|n=0,1,\ldots,N-1\}$ over all D days and each user has one and only one $\mathbb{C}$.

Definition 6 (Temporal Correlation): A user's temporal correlation of length I, denoted by $\mathbb{C}$, depicts the relation between two locations at current time slot t_n and its following i-th time slots t_{n+i}, and its definition is as below:

$$\mathbb{C}=\left\{C_i|i=0,1,\ldots,I-1\right\},$$

with C_i being the ensemble average of the locations pair at time slots (t_{n+i},t_n), denoted as (x_{n+i},x_n), over all days,

$$\begin{aligned}C_i&=\mathbf{E}_d\left\{x_{n+i}^{(d)}x_n^{(d)}\right\}\\&=\int x_{n+i}x_n dPr\left(x_{n+i}|x_n\right)dPr\left(x_n\right),\end{aligned}$$

where $d\in\{1,2,\ldots,D\}$.

From **Lemma 2** in **Appendix B**, it can be shown that, the Fourier coefficients of $\mathbb{C}$ is the variance of the Fourier

coefficients of a user's daily trajectories over all days, *i.e.*, $\mathcal{S}_k = \mathbf{Var}_d\left\{\mathcal{F}_k^{(d)}\right\} = 2\sigma^2_{\mathcal{S}_k}$,

$$C_i = \sum_{k=0}^{K-1}\left(2\sigma^2_{\mathcal{S}_k}\right) W_{i,k}, \tag{4}$$

where $k = 0, 1, \ldots, K-1$ and $d = 1, 2, \ldots, D$.

If the trajectories set follows the ergodic process, the ensemble average of the temporal correlation is equivalent to an autocorrelation function within a single daily trajectory,

$$C_i = \frac{1}{N-i+1}\sum_{n=0}^{N-i} x_{n+i}x_n.$$

C. THE DFT OF THE TEMPORAL CORRELATION

The DFT coefficients of a user's temporal correlation $\mathbb{C}$, denoted as $\left\{\mathcal{S}_k \middle| k = 0, 1, \ldots, K-1\right\}$, is obtained from Eq. (4),

$$\mathcal{S}_k = 2\sigma^2_{\mathcal{S}_k}. \tag{5}$$

For ergodic trajectories, we also have,

$$\mathcal{S}_k = |\mathcal{F}_k|^2, \tag{6}$$

where $\mathcal{F}_k$ is the k-th DFT coefficient of the user's daily trajectories given in **Definition 3**.

V. THE FGS-PUFFERFISH PRIVACY MECHANISM

In this section, we present a noise mechanism based on the geometric sum of the Fourier coefficients and optimize the data utility for a given privacy budget.

A. THE SETTING

1) TEMPORAL CORRELATION SECRET

For our problem, the Pufferfish secrets set consists of temporal correlations of all user given in **Definition 6**, denoted by $\mathcal{C}$, and it is expressed as follows,

$$\mathcal{C} = \left\{\mathbb{C}^{(u_m)} \middle| \, m = 1, 2, \ldots, M\right\}.$$

2) CORRELATION SECRETS PAIRS

A correlation secrets pair consists of two temporal correlations of any two users u and v in the same database, denoted as $\mathcal{Q}$,

$$\mathcal{Q} = \left\{\left(\mathbb{C}^{(u)}, \mathbb{C}^{(v)}\right) \middle| \, u, v \in \{u_1, u_2, \cdots, u_M\}, u \neq v\right\}.$$

3) THE ADVERSARY'S BELIEF DISTRIBUTION

The adversary's belief distribution θ is the probability of the temporal correlation $\mathbb{C}$,

$$\Theta = \{\theta : Pr\,(\mathbb{C})\}.$$

B. NOISE MECHANISM

From Eq. (5) and Eq. (6), we know that the Fourier coefficients of the temporal correlation $\mathbb{C}$ are closely related to those of the daily trajectory. Thus it is natural to add noise to the Fourier coefficients of the trajectory. Actually, Rastogi and Nath [11] proposed a similar noise mechanism, *i.e.*, the so-called FPA approach, based on the noisy Fourier coefficients mechanism. Unfortunately, the FPA approach considers neither the temporal correlation issue, nor the data utility optimization problem.

In this paper, we propose to optimize the Fourier coefficients noise for the problem of temporal correlation privacy. First, we add the noise in the Fourier coefficients $\left\{\mathcal{F}'_k \middle| k = 0, 1, \ldots, K-1\right\}$ shown in **Definition 3**; then, we obtain the generated noisy daily trajectory $\{x'_n | n = 0, 1, \ldots, N-1\}$ according to the IDFT given in **Definition 4**, with the noisy Fourier coefficients $\left\{\mathcal{F}'_k \middle| k = 0, 1, \ldots, K-1\right\}$; and finally, the noisy temporal correlation $\mathbb{C}'$ is obtained from the noisy daily trajectory $\{x'_n | n = 0, 1, \ldots, N-1\}$ according to **Definition 6**.

1) THE FOURIER COEFFICIENTS NOISE MECHANISM

In the Fourier coefficients noise mechanism, we add the noise into the Fourier coefficient, and we have,

$$\mathcal{F}'_k = \mathcal{F}_k + \delta\mathcal{F}_k, \quad (k = 0, 1, \ldots, K-1). \tag{7}$$

The Fourier coefficients noise mechanism does not require specific probability distribution of $\left\{\delta\mathcal{F}_k \middle| k = 0, 1, \cdots, K-1\right\}$, as long as they are mutual independent. However, to have minimum K to achieve Laplace distribution of the geometric sum, a natural choice is the Laplace distribution for both the real part and imaginary part of the noisy Fourier coefficients, *i.e.*, $\left\{\left(\delta\mathcal{F}^r_k, \delta\mathcal{F}^i_k\right) \middle| k = 0, 1, \ldots, K-1\right\}$,

$$Pr\left(\delta\mathcal{F}^r_k\right) = \exp\left(-\frac{\left|\delta\mathcal{F}^r_k\right|}{b_{\mathcal{F}_k}}\right) = Lap\left(\delta\mathcal{F}^r_k; 0, b_{\mathcal{F}_k}\right), \tag{8}$$

where we have only shown the set of real parts due to the similarity and assumed that both the set of real parts $\left\{\delta\mathcal{F}^r_k \middle| k = 0, 1, \ldots, K-1\right\}$ and the set of imaginary parts $\left\{\delta\mathcal{F}^i_k \middle| k = 0, 1, \ldots, K-1\right\}$ have the same Laplace noise scale parameter $b_{\mathcal{F}_k}$. Also, we have the following Laplace properties,

$$\begin{cases} \mathbf{E}_{\delta\mathcal{F}_k}\{\delta\mathcal{F}_k\} = 0, \\ \mathbf{Var}_{\delta\mathcal{F}_k}\{\delta\mathcal{F}_k\} = 2\,b^2_{\mathcal{F}_k}, \\ \mathbf{E}_{\delta\mathcal{F}_k}\left\{(\delta\mathcal{F}_k)^4\right\} = 24\,b^4_{\mathcal{F}_k}, \\ \mathbf{Var}_{\delta\mathcal{F}_k}\left\{(\delta\mathcal{F}_k)^2\right\} = 20\,b^4_{\mathcal{F}_k}. \end{cases}$$

2) LAPLACE LOCATION NOISE

After the addition of Fourier coefficients noise, the noisy location becomes,

$$x'_n = \sum_{k=0}^{K-1}\mathcal{F}'_k W_{n,k} = x_n + \delta x_n,$$

with

$$\delta x_n = \sum_{k=0}^{K-1} \delta \mathcal{F}_k W_{n,k}.$$

Theorem 2: The noisy location via the Fourier coefficients noise mechanism has a Laplace probability distribution given by,

$$Pr(x'_n) = \prod_{n=0}^{N-1} \frac{1}{2b_x} \exp\left(\frac{-|x'_n - x_n|}{b_x}\right),$$

with

$$b_x = \sqrt{\frac{1}{2}\sum_{k=0}^{K-1}(1-p)^k \left(2\, b^2_{\mathcal{F}_k}\right)}.$$

Proof: With the help of **Theorem 1**, **Lemma 1** in **Appendix A** shows that each noisy location x'_n follows the Laplace distribution with the mean of $\mathbf{E}\{x'_n\} = x_n$ and the same scale parameter $b_x = b_{x_n}$. So the noisy daily trajectory $\{x'_n | n = 0, 1, \ldots, N-1\}$ has a joint Laplace distribution and **Theorem 2** is proved.

3) LAPLACE TEMPORAL CORRELATION NOISE

Theorem 3: The noisy temporal correlation $\mathbb{C}'$ has a joint Laplace probability distribution given by,

$$Pr(\mathbb{C}') = \prod_{i=0}^{I-1} \frac{1}{2b_{C_i}} \exp\left(\frac{-|\delta C'_i - \delta \bar{C}_i|}{b_{C_i}}\right), \tag{9}$$

and

$$\begin{aligned}
\delta C'_i &= C'_i - \bar{C}_i,\\
\bar{C}_i &= \sum_{k=0}^{\infty}(1-p)^k \left(\sigma^2_{2\mathcal{S}_k}\right) W_{i,k},\\
\delta \bar{C}_i &= \sum_{k=0}^{\infty}(1-p)^k \left(2b^2_{\mathcal{F}_k}\right) W_{i,k},\\
b_{C_i} &= \sqrt{\sum_{k=0}^{\infty}(1-p)^k \left(12\, b^4_{\mathcal{F}_k}\right) - \frac{\left(\delta \bar{C}_i\right)^2}{2}}.
\end{aligned}$$

where $\{\bar{C}_i | i = 0, 1, \ldots, I-1\}$ is a set of the mean of the temporal correlation $\mathbb{C}$ before noise is added; and $\{\delta \bar{C}_i | i = 0, 1, \ldots, I-1\}$ is a set of the noise induced change of the mean of temporal correlation $\mathbb{C}$.

Proof: From **Lemma 3** (see **Appendix C**), we know that each C_i follows the Laplace distribution with the scale parameter b_{C_i}. Thus the temporal correlation $\mathbb{C} = \{C_i | i = 0, 1, \ldots, I-1\}$ has a joint probability distribution and **Theorem 3** is proved.

C. THE CONSTRAINED OPTIMIZATION OF THE FGS-PUFFERFISH PRIVACY MECHANISM

In this section, we propose to achieve the optimal data utility for a given privacy budget, *i.e.*, the constrained optimization problem. We resort to the LM method for such constrained optimization of the FGS-Pufferfish privacy mechanism.

Before the constrained optimization problem, let's quantify the data utility. Here we consider two utilities, including the location utility and the correlation utility. This is because, on one hand, the location utility is important for location based services such as location recommendation; and on the other hand, the correlation utility is important for applications such as location forecasting.

For the location utility, we want the average noisy location deviates from its raw location as small as possible and is defined as its variance given below,

Definition 7 (Location Utility): The location utility, denoted by $\mathcal{U}_L$, is the average variance of noisy locations $\{x'_n | n = 0, 1, \ldots, N-1\}$ after the noise is added. We have,

$$\mathcal{U}_L = \mathbf{E}_n\left\{\sigma^2_{x_n}\right\} = \sum_{k=0}^{\infty}(1-p)^k \left(2\, b^2_{\mathcal{F}_k}\right).$$

For the correlation utility, we want the mean of the noisy temporal correlations $\bar{C}_i$ deviates from its raw value C_i as small as possible, which is defined as follows,

Definition 8 (Correlation Utility): The correlation utility, denoted by $\mathcal{U}_{\mathbb{C}}$, is the average of the deviation of the noisy correlation $\bar{C}_i$ from its raw value C_i, denoted as $\delta \bar{C}_i = \bar{C}_i - C_i$, over its correlation length of I. From **Theorem 3**, we have,

$$\begin{aligned}
\mathcal{U}_{\mathbb{C}} &= \frac{1}{I}\sum_{i=0}^{I-1} \delta \bar{C}_i\\
&= \sum_{k=0}^{\infty}(1-p)^k \left(\mathcal{W}_k b^2_{\mathcal{F}_k}\right),
\end{aligned}$$

where $\mathcal{W}_k = \frac{1}{I}\sum_{i=0}^{I-1} 2\, W_{i,k}$.

Now let's define the data utility through the location utility and the correlation utility. And its definition is as follow.

Definition 9 (Data Utility): The data utility, denoted by $\mathcal{U}$, is the weighted sum of both the location utility $\mathcal{U}_L$ and the correlation utility $\mathcal{U}_{\mathbb{C}}$, and its definition is as below,

$$\begin{aligned}
\mathcal{U} &= w\mathcal{U}_L + (1-w)\mathcal{U}_{\mathbb{C}}\\
&= \sum_{k=0}^{\infty}(1-p)^k \left(\mathcal{W}'_k b^2_{\mathcal{F}_k}\right),
\end{aligned}$$

where $\mathcal{W}'_k = 2\, w + \frac{1-w}{I}\sum_{i=0}^{I-1} 2\, W_{i,k}$, with $w \in [0, 1]$ being the weight of the location utility $\mathcal{U}_L$.

It is well-known that the privacy budget ε is inversely proportional to the Laplace noise scale parameter, *i.e.*, the larger the Laplace noise scale parameter, the smaller the privacy budget ε or the better the privacy protection. For such reason, let's look at the Laplace noise scale parameters set $b_{\mathbb{C}} = \{b_{C_i} | i = 0, 1, \ldots, I-1\}$. According to **Theorem 3**, $b_{\mathbb{C}}$ has minimum value at b_{C_0}. So we only need to consider

b_{C_0} when dealing with the constrained optimization problem,

$$b_{C_0} = \sum_{k=0}^{\infty}(1-p)^k \left(12\, b_{\mathcal{F}_k}^4\right) - 2\left(\sum_{k=0}^{\infty}(1-p)^k b_{\mathcal{F}_k}^2\right)^2.$$

Now we can simplify the LM method in Eq. (3) for the constrained optimization problem of the FGS-Pufferfish privacy mechanism, with the help of **Theorem 3**,

$$\begin{cases} \min\left\{\mathcal{U} = \sum_{k=0}^{\infty}(1-p)^k \left(\mathcal{W}_k' b_{\mathcal{F}_k}^2\right)\right\}, \\ b_{C_0}^2 = \sum_{k=0}^{\infty}(1-p)^k \left(12\, b_{\mathcal{F}_k}^4\right) - 2\left(\sum_{k=0}^{\infty}(1-p)^k b_{\mathcal{F}_k}^2\right)^2, \end{cases} \tag{10}$$

with the Lagrangian of

$$\mathcal{L} = \sum_{k=0}^{\infty}(1-p)^k \left(\mathcal{W}_k' b_{\mathcal{F}_k}^2\right) - \lambda\left\{\sum_{k=0}^{\infty}(1-p)^k \left(12\, b_{\mathcal{F}_k}^4\right) - 2\left(\sum_{k=0}^{\infty}(1-p)^k b_{\mathcal{F}_k}^2\right)^2 - b_{C_0}^2\right\}.$$

From **Lemma 4** in **Appendix D**, the optimal Fourier noise sale parameters $b_{\mathcal{F}_k}^*$ is given below,

where $\overline{\overline{I}}$ is the unit matrix, $\overline{W'} = \left\{\mathcal{W}_k' | k = 0, 1, \cdots, K-1\right\}$, and

$$\overline{\overline{Q}} = \begin{bmatrix} (1-p)^0 & (1-p)^1 & (1-p)^2 & (1-p)^3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \cdots \\ (1-p)^0 & (1-p)^1 & (1-p)^2 & (1-p)^3 & \cdots \end{bmatrix}.$$

VI. ANALYSIS

In this section, we give the analysis of the data utility, privacy and the adversary knowledge.

A. UTILITY

According to the Eq. (11), as shown at the top of the next page, and the data utility given in **Definition 9**, we have the optimal data utility as below,

$$\mathcal{U}^* = \sum_{k=0}^{\infty}(1-p)^k \left(\mathcal{W}_k' b_{\mathcal{F}_k}^{*2}\right).$$

B. PRIVACY

First, let's look at the measure of the temporal correlation for the FGS-Pufferfish privacy mechanism.

A temporal correlation measure, denoted by $M(\mathbb{C})$, is the joint probability distribution of the noisy temporal correlation $\mathbb{C}'$ after noise is added,

$$Pr(M(\mathbb{C})) = Pr(\mathbb{C}') = \prod_{i=0}^{I-1} Pr(C_i').$$

Theorem 4: The Fourier coefficients noise mechanism satisfies ε-Pufferfish privacy of the noisy temporal correlation $\mathbb{C}'$ for a given noise scale parameter $b_{\mathbb{C}}$,

$$\exp(-\varepsilon) \le \frac{Pr\left(\mathbb{C}' \middle| \mathbb{C}^{(u)}, b_{\mathbb{C}}\right)}{Pr\left(\mathbb{C}' \middle| \mathbb{C}^{(v)}, b_{\mathbb{C}}\right)} \le \exp(\varepsilon),$$

with $\varepsilon = \varepsilon_0 + \Delta$, where

$$\varepsilon_0 = \sup_{u,v}\left\{\sum_{i=0}^{I-1}\left(\frac{\left|C_i^{(u)} - C_i^{(v)}\right|}{b_{C_i}}\right)\right\},$$

$$\Delta = \sup_{u,v}\left\{\sum_{i=0}^{I-1}\left(\frac{\left|\delta\bar{C}_i^{(u)} - \delta\bar{C}_i^{(v)}\right|}{b_{C_i}}\right)\right\}.$$

Proof: From **Theorem 3**, we have,

$$\frac{Pr\left(\mathbb{C}' \middle| \mathbb{C}^{(u)}, b_{\mathbb{C}}\right)}{Pr\left(\mathbb{C}' \middle| \mathbb{C}^{(v)}, b_{\mathbb{C}}\right)} = \exp\left\{-\sum_{i=0}^{I-1}\left(\frac{\left|\delta C_i^{(u)} - \delta\bar{C}_i^{(u)}\right| - \left|\delta C_i^{(v)} - \delta\bar{C}_i^{(v)}\right|}{b_{C_i}}\right)\right\},$$

from which we have,

$$\min_{\mathbb{C}'}\left\{\frac{Pr\left(\mathbb{C}' \middle| \mathbb{C}^{(u)}, b_{\mathbb{C}}\right)}{Pr\left(\mathbb{C}' \middle| \mathbb{C}^{(v)}, b_{\mathbb{C}}\right)}\right\} = \exp\left\{-\sum_{i=0}^{I-1}\left(\frac{\left|C_i^{(u)} - C_i^{(v)} + \delta\bar{C}_i^{(u)} - \delta\bar{C}_i^{(v)}\right|}{b_{C_i}}\right)\right\},$$

$$\max_{\mathbb{C}'}\left\{\frac{Pr\left(\mathbb{C}' \middle| \mathbb{C}^{(u)}, b_{\mathbb{C}}\right)}{Pr\left(\mathbb{C}' \middle| \mathbb{C}^{(v)}, b_{\mathbb{C}}\right)}\right\} = \exp\left\{\sum_{i=0}^{I-1}\left(\frac{\left|C_i^{(u)} - C_i^{(v)} + \delta\bar{C}_i^{(v)} - \delta\bar{C}_i^{(v)}\right|}{b_{C_i}}\right)\right\}.$$

So the privacy budget ε is obtained by taking the maximum difference of the probability ratios for all pairs of trajectories $\left(\mathbb{C}^{(u)}, \mathbb{C}^{(v)}\right)$,

$$\begin{aligned} \varepsilon &= \sup_{u,v}\left\{-\ln\left(\min_{\mathbb{C}'}\left\{\frac{Pr\left(\mathbb{C}';\ \mathbb{C}^{(u)}, b_{\mathbb{C}}\right)}{Pr\left(\mathbb{C}';\ \mathbb{C}^{(v)}, b_{\mathbb{C}}\right)}\right\}\right)\right\} \\ &= \sup_{u,v}\left\{\ln\left(\max_{\mathbb{C}'}\left\{\frac{Pr\left(\mathbb{C}';\ \mathbb{C}^{(u)}, b_{\mathbb{C}}\right)}{Pr\left(\mathbb{C}';\ \mathbb{C}^{(v)}, b_{\mathbb{C}}\right)}\right\}\right)\right\} \\ &= \sup_{u,v}\left\{\sum_{i=0}^{I-1}\left(\frac{\left|C_i^{(u)} - C_i^{(v)} + \delta\bar{C}_i^{(u)} - \delta\bar{C}_i^{(v)}\right|}{b_{C_i}}\right)\right\} \\ &= \sup_{u,v}\left\{\sum_{i=0}^{I-1}\left(\frac{\left|C_i^{(u)} - C_i^{(v)}\right|}{b_{C_i}}\right) + \sum_{i=0}^{I-1}\left(\frac{\left|\delta\bar{C}_i^{(u)} - \delta\bar{C}_i^{(v)}\right|}{b_{C_i}}\right)\right\} \\ &= \varepsilon_0 + \Delta. \end{aligned}$$

and **Theorem 4** is proved.

$$b^*_{\mathcal{F}_k} = \sqrt{\frac{b_{C_0}\left[\left(6\bar{\bar{I}} - \bar{\bar{Q}}\right)^{-1}\frac{\overline{W'}}{2}\right]_k}{\sqrt{3\sum_{k=0}^{\infty}(1-p)^k\left[\left(6\bar{\bar{I}} - \bar{\bar{Q}}\right)^{-1}\overline{W'}\right]_k^2 - \frac{\left(\sum_{k=0}^{\infty}(1-p)^k\left[\left(6\bar{\bar{I}} - \bar{\bar{Q}}\right)^{-1}\overline{W'}\right]_k\right)^2}{2}}}}. \tag{11}$$

C. ADVERSARY KNOWLEDGE

Our goal is to prevent an adversary from mining a user's privacy through analyzing the user's temporal correlation based on the adversary's prior knowledge about the user. A good privacy mechanism requires that the adversary's posterior-to-prior knowledge gain should be close to unit 1. Mathematically, the adversary's knowledge gain is given by the ratio of the posterior probability $Pr(\mathbb{C}' \sim 0|\mathbb{C})$ to the prior probability $Pr(\mathbb{C}' \sim 0)$, given that the adversary has known the raw daily trajectory correlation of the user $\mathbb{C}^{(u)}$.

1) PRIOR KNOWLEDGE

Without any knowledge of the correlation of a user $\mathbb{C}^{(u)}$, the adversary might uniformly pick $\mathbb{C}$. From **Lemma** 3 (see **Appendix C**), the adversary's prior knowledge is given by,

$$Pr\left(\mathbb{C}'\right) = \int_{\mathbb{C}} \prod_{i=0}^{I-1} Pr\left(C'_i; \bar{\mathbb{C}}, b_{\mathbb{C}}\right) Pr(\mathbb{C}) d\mathbb{C}. \tag{12}$$

2) POSTERIOR KNOWLEDGE

Given that the adversary has known the user's temporal correlation $\mathbb{C}^{(u)}$, the adversary's posterior knowledge is given by,

$$Pr\left(\mathbb{C}'|\mathbb{C}^{(u)}\right) = \prod_{i=0}^{I-1} Pr\left(C'_i; \bar{\mathbb{C}}^{(u)}, b_{\mathbb{C}}\right). \tag{13}$$

3) POSTERIOR-OVER-PRIOR KNOWLEDGE GAIN

Now, let's look at the posterior-over-prior knowledge gain of an adversary.

Theorem 5: The posterior-over-prior knowledge gain of an adversary satisfies

$$\exp(-\varepsilon) \leq \frac{Pr(\mathbb{C}'|\mathbb{C})}{Pr(\mathbb{C}')} \leq \exp(\varepsilon).$$

Proof: From Eq. (12) and Eq. (13), we have,

$$\begin{aligned}
&\frac{Pr(\mathbb{C}')}{Pr(\mathbb{C}'|\mathbb{C}^{(u)})} \\
&= \frac{\int_{\mathbb{C}} \prod_{i=0}^{I-1} Pr\left(C'_i;\ \bar{\mathbb{C}}, b_{\mathbb{C}}|\mathbb{C}\right) Pr(\mathbb{C}) d\mathbb{C}}{\prod_{i=0}^{I-1} Pr\left(C'_i;\ \bar{\mathbb{C}}^{(u)}, b_{\mathbb{C}}|\mathbb{C}^{(u)}\right)} \\
&= \int_{\mathbb{C}} \left[\prod_{i=0}^{I-1} Pr(C'_i;\ \bar{\mathbb{C}}, b_{\mathbb{C}}) \Big/ \prod_{i=0}^{I-1} Pr(C'_i;\ \bar{\mathbb{C}}^{(u)}, b_{\mathbb{C}})\right] Pr(\mathbb{C}) d\mathbb{C}.
\end{aligned}$$

From **Theorem 4**, we have,

$$\exp(-\varepsilon) \leq \frac{Pr(\mathbb{C}'|\mathbb{C}^{(u)})}{Pr(\mathbb{C}')} \leq \exp(\varepsilon),$$

and **Theorem 5** is proved.

VII. THE PRIVACY-PRESERVING TRAJECTORIES RELEASING

In this section, we summarize our FGS-Pufferfish privacy mechanism and present the numerical recipe of the privacy-preserving daily trajectory releasing algorithm.

The basic reasoning of the FGS-Pufferfish privacy mechanism is as follows:

1) First, we define the constrained optimization problem of achieving a better data utility $\mathcal{U}^*$ for a given privacy budget given by the Laplace scale parameter of C_0, *i.e.*, b_{C_0}.
2) Next, we solve the constrained optimization problem via the LM method and obtain the optimal obtained Laplace scale parameter $\left\{b^*_{\mathcal{F}_k}|k = 0, 1, \ldots, K-1\right\}$ for the noisy Fourier coefficients as shown in Eq. (11).
3) Then, the FGS-Pufferfish privacy mechanism adds noise to the Fourier coefficients according to Eq. (7) and obtain the noisy Fourier coefficients, *i.e.*, $\left\{\mathcal{F}'_k|k = 0, 1, \ldots, K-1\right\}$.
4) At last, we obtain the sanitized daily trajectories $\left\{\left(x'_n, y'_n\right)|n = 0, 1, \ldots, N-1\right\}$ according to Eq. (2), with the noisy Fourier coefficients $\left\{\mathcal{F}'_k|k = 0, 1, \ldots, K-1\right\}$.

Based our proposed FGS-Pufferfish privacy mechanism, we design an algorithm to release temporally correlated trajectories in order to protect individuals' privacy. Because our goal is to protect the temporal correlation of a user's daily trajectory, we first calculate the Fourier coefficients $\{\mathcal{F}_k|k = 0, 1, \ldots, K-1\}$ of a daily trajectory $\{x_n|n = 0, 1, \ldots, N-1\}$, which is related to the Fourier coefficients $\{\mathcal{S}_k|k = 0, 1, \ldots, K-1\}$ of its temporal correlation $\mathbb{C}$. Then, to achieve the Laplace distribution of the noisy temporal correlation $\mathbb{C}'$ through the Fourier coefficients noise mechanism, *i.e.*, adding noise to K Fourier coefficients, with K following the geometric distribution. Furthermore, we obtain the optimal Laplace scale parameters $\left\{b^*_{\mathcal{F}_k}|k = 0, 1, \ldots, K-1\right\}$ for the noisy Fourier coefficients. Finally, we use IDFT to obtain the noisy locations of the sanitized daily trajectory. The details of the privacy-preserving trajectory releasing through the FGS-Pufferfish privacy mechanism is given in **Algorithm** 1.

Algorithm 1 FGS-Pufferfish Private Trajectory Releasing

Input: A raw daily trajectories set: $\mathcal{T} = \left\{\mathbb{T}^{(d,u_m)} \middle| m = 1, 2, \ldots, M; d = 1, 2, \ldots, D\right\}$, and a given Laplace scale parameter of C_0: b_{C_0}.

Output: A privacy-preserving trajectories set: $\mathcal{T}' = \left\{\mathbb{T}'^{(d,u_m)} \middle| m = 1, 2, \ldots, M; d = 1, 2, \ldots, D\right\}$.

for all $m \in \{1, 2, \ldots, M\}$ **do**
  $\mathbb{T}'^{(d,u_m)} = \emptyset$
  Calculate the ensemble average of $(\overline{X}_n, \overline{Y}_n)$.
  for all $d \in \{1, 2, \ldots, D\}$ **do**
    Obtain the modified daily trajectory of the d-th day: $\left\{(x_n, y_n) \middle| n = 0, 1, \ldots, N-1\right\}$, where $x_n = X_n - \overline{X}_n$ and $y_n = Y_n - \overline{Y}_n$ (only $\{x_n | n = 0, 1, \ldots, N-1\}$ is shown).
    Calculate the Fourier coefficients for the modified longitude: $\mathcal{F}_k = \sum_{i=0}^{N-1} x_n \exp\left(-j\frac{2\pi}{N}kn\right)$.
    Select an appropriate p and obtain K through the geometric distribution: $Pr(K) = (1-p)^{K-1}p$.
    Select the utility weight w and calculate $\mathcal{W}'_k$:

$$\mathcal{W}'_k = 2\,w + (1-w)\mathcal{W}_k = 2\,w + \frac{1-w}{I}\sum_{i=0}^{I-1} 2\,W_{i,k}.$$

    Calculate the optimal Laplace scale parameter for the noisy Fourier coefficients:

$$b^*_{\mathcal{F}_k} = \sqrt{\frac{b_{C_0}\left[\left(6\bar{l}-\bar{\bar{Q}}\right)^{-1}\frac{W'}{2}\right]_k}{\sqrt{3\sum_{k=0}^{\infty}(1-p)^k\left[\left(6\bar{l}-\bar{\bar{Q}}\right)^{-1}\overline{W'}\right]_k^2 - \frac{\left(\sum_{k=0}^{\infty}(1-p)^k\left[\left(6\bar{l}-\bar{\bar{Q}}\right)^{-1}\overline{W'}\right]_k\right)^2}{2}}}}.$$

    Obtain noisy Fourier coefficients $\mathcal{F}'_k = \mathcal{F}_k + \delta\mathcal{F}_k$:

$$Pr(\delta\mathcal{F}_k) = \exp\left(-\frac{|\delta\mathcal{F}_k|}{b^*_{\mathcal{F}_k}}\right).$$

    for all $n \in \{0, 1, \ldots, N-1\}$ **do**
      Obtain the noisy location x'_n through IDFT: $x'_n = \sum_{k=0}^{K-1} \mathcal{F}'_k \exp\left(j\frac{2\pi}{N}kn\right)$.
    end for
    Obtain the noisy modified daily trajectory:

$$\left\{\left(x'_n, y'_n\right) \middle| n = 0, 1, \ldots, N-1\right\}.$$

    $\mathbb{T}'^{(d,u_m)} = \mathbb{T}'^{(d,u_m)} \cup \left\{\left(x'_n + \bar{X}_n, y'_n + \bar{Y}_n\right) \middle| n = 0, 1, \ldots, N-1\right\}$.
  end for
end for
return $\mathcal{T}' = \left\{\mathbb{T}'^{(d,u_m)} \middle| m = 1, 2, \cdots, M; d = 1, 2, \cdots, D\right\}$.

VIII. EVALUATION

A. EVALUATION SETTING

1) CONFIGURATIONS

We implement our simulations with Python 2.7 on a laptop with Intel Core i7-6500U, 2.59GHz, Windows 10 system equipped with 8GB main memory. The trajectories' temporal correlations are evaluated at time intervals of $\left\{t_i = i\tau \middle| i = 0, 1, \ldots, I-1\right\}$, with τ being a constant, as required by the DFT and IDFT. Also, all simulations are for the data utility $\mathcal{U}$ that consists of $w = 90\%$ location utility $\mathcal{U}_L$ and $1 - w = 10\%$ correlation utility $\mathcal{U}_{\mathbb{C}}$. What's more, due to similarity, only the result for the location utility $\mathcal{U}_L$ is shown.

2) DATA

We compare our FGS-Pufferfish privacy mechanism to the CTS approach [3] on both simulated and real-life data [32]. Because the real-life data is not sampled at a constant time interval τ, we interpolate the raw data in order that the obtained data has a constant time interval τ.

a: SIMULATED DATA

The simulated data is consist of temporally correlated simulated trajectories which are generated through updating the current location x_n at the n-th time slot by keeping only a fraction of the last location x_{n-1}, while adding a random location change g that follows the Gaussian distribution, *i.e.*, $x_n = C_1 x_{n-1} + g\sqrt{1 - C_1^2}$, with $C_1 = \exp^{-1/\tau_c}$ being the temporal correlation at $i = 1$ and τ_c is the correlation time constant. The generated daily trajectory has an exponential temporal correlation of $C_i = \exp^{-i/\tau_c}$. This data contains 9 trajectories, and each daily trajectory consists of 32 time slots with a constant time interval. For both our FGS-Pufferfish privacy mechanism and the CTS approach, 1,000,000 noise realizations are used for statistics.

b: REAL DATA

The real data is collected by Yonsei University in Seoul of Korea. It contains nine users' mobility trajectories in 62 days. A daily trajectory in this dataset consists of longitudes x_n, latitudes y_n and timestamps t. After an interpolation with a constant time interval, each user's daily trajectory has 48 times slots. Similar to the simulated dataset, we used 1,000,000 noise realizations for statistics.

3) METRICS

During our evaluations, firstly, we evaluate the probability distribution for noisy temporal correlations, due to similarity reason, we only show the result of $Pr(C'_0)$. Then, we focus on other two metrics of performance: privacy and utility.

In the previous analysis, we have proved that our mechanism satisfies ε-Pufferfish privacy. Hence, we adopt the privacy budget ε to evaluate the privacy. For the privacy budget ε, we use the joint probability for all coefficients $\left\{C'_i \middle| i = 0, 1, \ldots, I-1\right\}$ of the correlation function $\mathbb{C}'$, *i.e.*, $Pr(\mathbb{C}')$ in Eq. (9).

Furthermore, for the location utility evaluation, we adopt the utility of the average location standard deviation of a daily trajectory $\mathcal{U}_L = \overline{\sigma}_L$ in **Definition 9** to measure the data utility.

B. SIMULATED DATA

We first generate correlated trajectories with an exponential temporal correlation. Then we evaluate the probability distribution of the noisy temporal correlation $Pr(C_0')$ as shown in Figure 2. It is clear that the distribution $Pr(C_0')$ of our proposed FGS-Pufferfish privacy mechanism is a symmetric Laplace distribution, agreeing with **Theorem 3**, while the probability distribution $Pr(C_0')$ of the CTS approach is neither symmetric and nor a Laplace distribution.

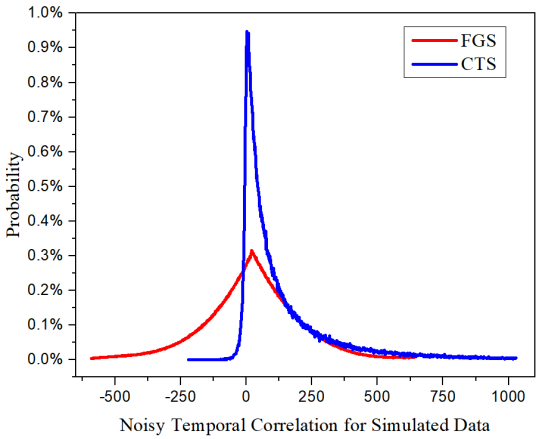


FIGURE 2. Temporal correlation distribution $Pr\left(C_0'\right)$ for simulated data.

1) PRIVACY EVALUATION FOR SIMULATED DATA

Now, let's see the privacy budget ε. As we know, the smaller the privacy budget ε, the better the privacy. The experimental results are shown in Figure 3, from which it is clear that our FGS-Pufferfish privacy mechanism does achieve better privacy than the CTS approach, for given location utilities $\mathcal{U} = \overline{\sigma}_L$.

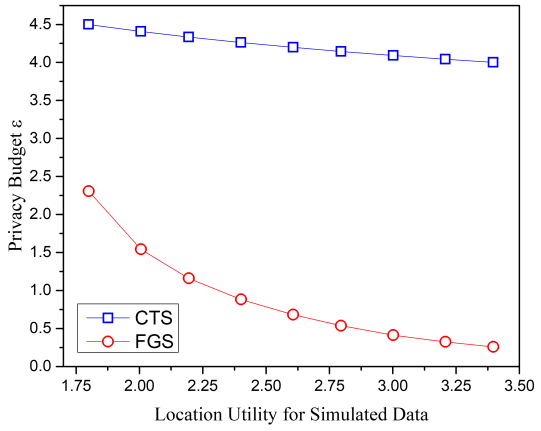


FIGURE 3. Privacy budget ε evaluation for simulated data.

2) UTILITY EVALUATION FOR SIMULATED DATA

Now, let's look at the location utility $\mathcal{U}_L$ under the same privacy budget ε. As shown in Figure 4, the location utility $\mathcal{U}_L$ of our FGS-Pufferfish privacy mechanism is better than that of the CTS approach. Also, we can see that the smaller the privacy budget ε, the better the location utility $\mathcal{U}_L$.

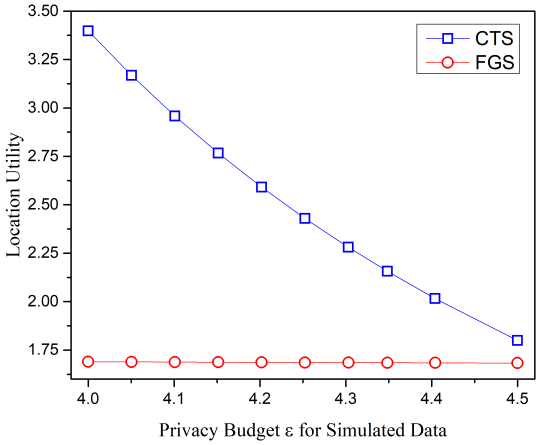


FIGURE 4. Utility $\mathcal{U}$ evaluation for simulated data.

C. REAL DATA

For real-life data, we evaluate the longitude coordinate x_n and the latitude coordinate y_n independently. Similar to the simulated data, we first evaluate the probability distribution $Pr\left(C_{x,0}'\right)$ and $Pr\left(C_{y,0}'\right)$ of the temporal correlations after noise is added, and the results as shown in Figure 5 for $Pr\left(C_{x,0}'\right)$ and Figure 6 for $Pr\left(C_{y,0}'\right)$. Similar to the simulated data, both $Pr\left(C_{x,0}'\right)$ and $Pr\left(C_{y,0}'\right)$ of our proposed FGS-Pufferfish privacy mechanism are symmetric Laplace distributions, agreeing with **Theorem 3**, while those of the CTS approach are not.

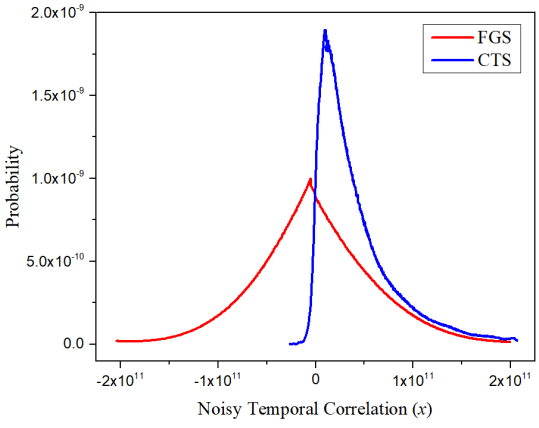


FIGURE 5. Temporal correlation distribution $Pr\left(C_{x,0}'\right)$ for longitudes.

1) PRIVACY EVALUATION FOR REAL DATA

Now let's look at the privacy budgets ε_x and ε_y. As shown in the Figure 7 and Figure 8, the privacy budgets ε_x and ε_y of our FGS-Pufferfish privacy mechanism are smaller than

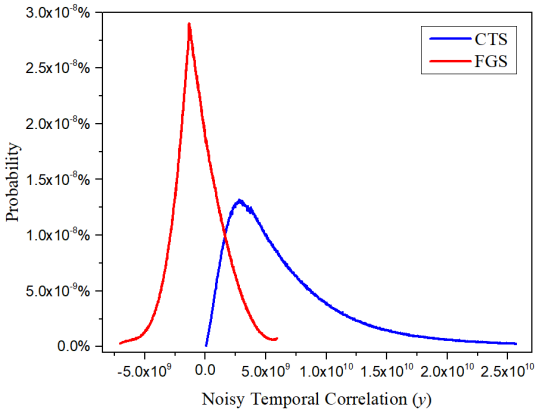


FIGURE 6. Temporal correlation distribution $Pr\left(C'_{y,0}\right)$ for latitudes.

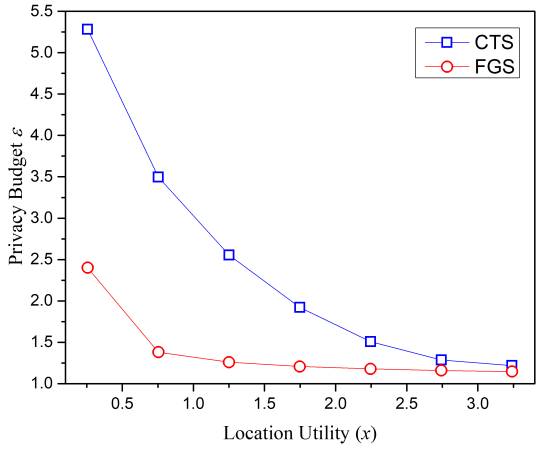


FIGURE 7. Privacy budget ε_x evaluation for longitudes.

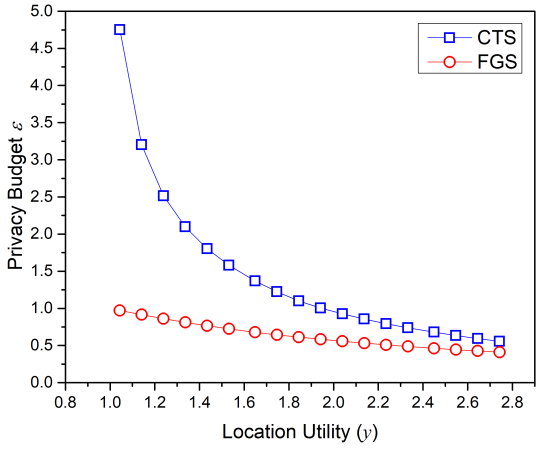


FIGURE 8. Privacy budget evaluation ε_y for latitudes.

those of the CTS approach for the same given data utilities $\mathcal{U}_{x_x}$ and $\mathcal{U}_{x_y}$. Also, both approaches show that the privacy is better for a worse data utility, and vice versa.

2) UTILITY EVALUATION FOR REAL DATA

Similar to the simulated data, we have also evaluated the location utilities $\mathcal{U}_{x_x}$ and $\mathcal{U}_{x_y}$ under the same privacy budget ε_x and ε_y, which are shown in Figure 9 and Figure 10. Again, it is clear that our FGS-Pufferfish privacy mechanism achieves better location utility for the same given privacy budgets ε_x and ε_y.

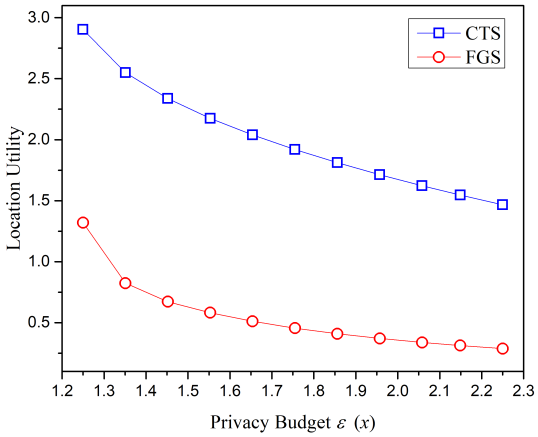


FIGURE 9. Location utility $\mathcal{U}_x$ evaluation for longitudes.

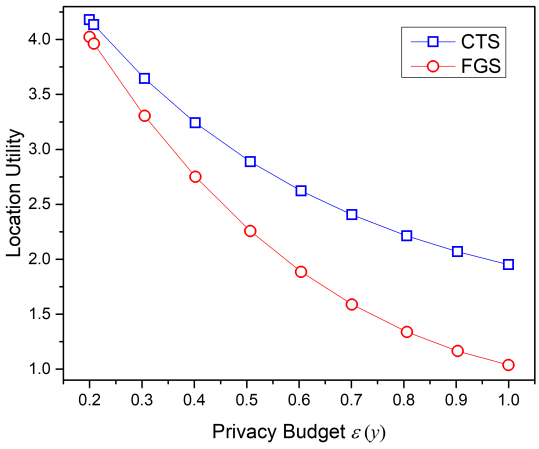


FIGURE 10. Location utility $\mathcal{U}_y$ evaluation for latitudes.

IX. CONCLUSION

We propose a Laplace noise mechanism based on the noisy Fourier coefficients' geometric sum, satisfying Pufferfish privacy, *i.e.*, the FGS-Pufferfish privacy mechanism, to protect the temporal correlation of a user's daily trajectories. The optimal noisy Fourier coefficients are obtained by solving the constrained optimization problem via the LM method to achieves a better data utility for a given privacy budget. Experiments with both simulated and real-life data show that our FGS-Pufferfish privacy mechanism achieves better data utility and privacy compared to the existing approach. Although we only deal with daily trajectories with a constant time interval, our proposed mechanism can be readily modified for time-series data with irregular time intervals. At last, the geometric sum can be combined with other decomposition methods such as Principal Components

Analysis (PCA) and wavelets methods for better performance of specific time-series applications.

APPENDIX A
NOISY LOCATION PROBABILITY DISTRIBUTION

Lemma 1: After noise is added to the Fourier coefficients, the noisy location's Laplace distribution is given by

$$Pr(x'_n) = Lap\left(x'_n; x_n, b_{x_n}\right),$$

with

$$b_{x_n} = \sqrt{\frac{1}{2}\sum_{k=0}^{\infty}(1-p)^k\left(2\,b^2_{\mathcal{F}_k}\right)}, \quad (n = 0, 1, \ldots, N-1).$$

Proof: From **Theorem 1**, we know that x'_n follows the Laplace distribution, and its mean $\mathbf{E}_{x'_n}$ and Laplace scale parameter b_{x_n} are given by,

$$\mathbf{E}_{x'_n}\left\{x'_n\right\} = x_n, \quad b_{x_n} = \frac{\sigma_{x_n}}{\sqrt{2}}.$$

Now let's calculate the standard deviation σ_{x_n},

$$\begin{aligned}
\sigma^2_{x_n} &= \mathbf{E}_K\left\{\sum_{k=0}^{K-1}\sum_{k'=0}^{K-1}\mathbf{E}_{\delta\mathcal{F}_k}\left\{\delta\mathcal{F}_k\delta\mathcal{F}^*_{k'}\right\}W^{-1}_{n,k}W_{n,k'}\right\}\\
&= \mathbf{E}_K\left\{\sum_{k=0}^{K-1}\left(2\,b^2_{\mathcal{F}_k}\right)\right\} = \sum_{k=0}^{\infty}\left(\sum_{K\geq k}^{\infty}Pr(K)\right)\left(2\,b^2_{\mathcal{F}_k}\right)\\
&= \sum_{k=0}^{\infty}\left(1-\sum_{K=0}^{K<k}Pr(K)\right)\left(2\,b^2_{\mathcal{F}_k}\right)\\
&= \sum_{k=0}^{\infty}(1-p)^k\left(2\,b^2_{\mathcal{F}_k}\right),
\end{aligned}$$

from which **Lemma 1** is proved.

APPENDIX B
TEMPORAL CORRELATION OF A USER

Lemma 2: The temporal correlation of a user before noise is added, denoted as $\mathbb{C}$, is the Fourier transform of the variance of the Fourier coefficients of the user's daily trajectories over all D days, *i.e.*, $\mathcal{S}_k = \mathbf{Var}_d\left\{\mathcal{F}^{(d)}_k\right\} = 2\sigma^2_{\mathcal{S}_k}$, $(k = 0, 1, \ldots, K-1; d = 1, 2, \ldots, D)$,

$$C_i = \sum_{k=0}^{K-1}\left(2\sigma^2_{\mathcal{S}_k}\right)W_{i,k}.$$

Proof: From Eq. (2), we have,

$$\begin{aligned}
C_i &= \mathbf{E}_d\left\{x^{(d)}_{n+i}x^{(d)}_n\right\}\\
&= \mathbf{E}_{\mathcal{F}_k}\left\{\sum_{k=0}^{K-1}\mathcal{F}_kW_{i+n,k}\sum_{k'=0}^{K-1}\mathcal{F}^*_{k'}W^{-1}_{n,k}\text{âĂŸ}\right\}\\
&= \sum_{k=0}^{K-1}W_{i+n,k}\sum_{k'=0}^{K-1}W^{-1}_{n,k'}\mathbf{E}_{\mathcal{F}_k}\left\{\mathcal{F}_k\mathcal{F}^*_{k'}\right\}\\
&= \sum_{k=0}^{K-1}W_{i+n,k}\sum_{k'=0}^{K-1}W^{-1}_{n,k'}\mathbf{E}_{\mathcal{F}_k}\left\{|\mathcal{F}_k|^2\right\}\delta(k-k')\\
&= \sum_{k=0}^{K-1}\left(2\sigma^2_{\mathcal{S}_k}\right)W_{i,k}.
\end{aligned}$$

where we have used the property of the Chi-Square χ^2_2 distribution in Eq. (1) and thus **Lemma 2** is proved.

APPENDIX C
NOISY TEMPORAL CORRELATION DISTRIBUTION

Similar to the noisy location distribution, we can obtain the noisy temporal correlation distribution according to the geometric sum given in **Theorem 1**.

Lemma 3: The noisy temporal correlation coefficient C'_i follows the Laplace distribution,

$$Pr(C'_i) = Lap\left(C'_i;\ \bar{C}'_i, b_{C_i}\right),$$

with

$$\begin{aligned}
\bar{C}'_i &= 2\sum_{k=0}^{\infty}(1-p)^k\left(\sigma^2_{\mathcal{S}_k} + b^2_{\mathcal{F}_k}\right)W_{i,k},\\
b_{C_i} &= \sqrt{\sum_{k=0}^{\infty}(1-p)^k\left(12\,b^4_{\mathcal{F}_k}\right) - \frac{\left(\delta\bar{C}_i\right)^2}{2}}.
\end{aligned}$$

Proof: Similar to **Lemma 2**, the noisy temporal correlation $\left\{C'_i|i = 0, 1, \cdots, I-1\right\}$ after the Fourier coefficients noise mechanism can be obtained,

$$\begin{aligned}
C'_i &= \sum_{k=0}^{K-1}\mathbf{E}_{\mathcal{F}_k}\left\{|\mathcal{F}'_k|^2\right\}W_{i,k}\\
&= \sum_{k=0}^{K-1}\left(\mathbf{E}_{\mathcal{F}_k}\left\{|\mathcal{F}_k|^2\right\} + |\delta\mathcal{F}_k|^2\right)W_{i,k}\\
&= \sum_{k=0}^{K-1}\left(2\sigma^2_{\mathcal{S}_k} + |\delta\mathcal{F}_k|^2\right)W_{i,k}. \qquad (14)
\end{aligned}$$

We can see that C'_i in Eq. (14) is in the form of geometric sum. From **Theorem 1**, we know that $\left\{C'_i|i = 0, 1, \ldots, I-1\right\}$ follows Laplace distribution. With the help of the properties of the Laplace Fourier coefficients noise $\left\{\delta\mathcal{F}_k|k = 0, 1, \ldots, K-1\right\}$ given in Eq. (8), the mean of the noisy temporal correlation coefficients $\left\{C'_i|i = 0, 1, \ldots, I-1\right\}$ after the Fourier coefficients noise mechanism is given by,

$$\begin{aligned}
\bar{C}'_i &= \mathbf{E}_d\left\{x'^{(d)}_{n+i}x'^{(d)}_n\right\}\\
&= \mathbf{E}_{(K,\delta\mathcal{F}_k)}\left\{\sum_{k=0}^{K-1}\left(2\sigma^2_{\mathcal{S}_k} + |\delta\mathcal{F}_k|^2\right)W_{i,k}\right\}\\
&= \bar{C}_i + \delta\bar{C}_i,
\end{aligned}$$

$$b^*_{\mathcal{F}_k} = \sqrt{\frac{b_{C_0}\left[\left(6\bar{\bar{I}} - \bar{\bar{Q}}\right)^{-1}\frac{\overline{W'}}{2}\right]_k}{\sqrt{3\sum_{k=0}^{\infty}(1-p)^k\left[\left(6\bar{\bar{I}} - \bar{\bar{Q}}\right)^{-1}\overline{W'}\right]_k^2 - \frac{\left(\sum_{k=0}^{\infty}(1-p)^k\left[\left(6\bar{\bar{I}} - \bar{\bar{Q}}\right)^{-1}\overline{W'}\right]_k\right)^2}{2}}}}. \tag{15}$$

$$\lambda^* = \frac{\sqrt{3\sum_{k=0}^{\infty}(1-p)^k\left[\left(6\bar{\bar{I}} - \bar{\bar{Q}}\right)^{-1}\overline{W'}\right]_k^2 - \frac{\left(\sum_{k=0}^{\infty}(1-p)^k\left[\left(6\bar{\bar{I}} - \bar{\bar{Q}}\right)^{-1}\overline{W'}\right]_k\right)^2}{2}}}{2b_{C_0}}. \tag{17}$$

where $\bar{C}_i$ is the mean of the temporal correlation before noise is added; and $\delta\bar{C}_i$ is the noise induced change of the mean of the temporal correlation $\mathbb{C}$,

$$\bar{C}_i = \mathbf{E}_{(K,\delta\mathcal{F}_k)}\left\{\sum_{k=0}^{K-1}\left(2\sigma^2_{\mathcal{S}_k}\right)W_{i,k}\right\},$$
$$\delta\bar{C}_i = \mathbf{E}_{(K,\delta\mathcal{F}_k)}\left\{\sum_{k=0}^{K-1}|\delta\mathcal{F}_k|^2 W_{i,k}\right\}.$$

Now let's calculate $\bar{C}_i$ and $\delta\bar{C}_i$,

$$\bar{C}'_i = \mathbf{E}_K\left\{\sum_{k=0}^{K-1}\left(2\sigma^2_{\mathcal{S}_k}\right)W_{i,k}\right\} = 2\sum_{k=0}^{\infty}(1-p)^k\left(\sigma^2_{\mathcal{S}_k}\right)W_{i,k}.$$

Similarly, we have,

$$\delta\bar{C}_i = \sum_{k=0}^{\infty}(1-p)^k\left(2\,b^2_{\mathcal{F}_k}W_{i,k}\right).$$

We can also obtain the variance of C'_i $(i = 0, 1, \ldots, I-1)$ as follows,

$$\begin{aligned}
&\mathbf{Var}_{(K,\delta\mathcal{F}_k)}\left\{C'_i\right\}\\
&= \mathbf{Var}_{(K,\delta\mathcal{F}_k)}\left\{\sum_{k=0}^{K-1}\left(|\delta\mathcal{F}_k|^2\right)W_{i,k}\right\}\\
&= \mathbf{E}\left\{\sum_{k=0}^{K-1}\left(|\delta\mathcal{F}_k|^2\right)W_{i,k}\sum_{k'=0}^{K-1}\left(|\delta\mathcal{F}_{k'}|^2\right)W^{-1}_{i,k'}\right\} - \left(\delta\bar{C}_i\right)^2\\
&= \mathbf{E}_K\left\{\sum_{k=0}^{K-1}\mathbf{E}_{\delta\mathcal{F}_k}\left\{|\delta\mathcal{F}_k|^2\right\}\right\} - \left(\delta\bar{C}_i\right)^2\\
&= \sum_{k=0}^{\infty}(1-p)^k\left(24\,b^4_{\mathcal{F}_k}\right) - \left(\delta\bar{C}_i\right)^2 = 2b^2_{C_i}.
\end{aligned}$$

where we have used the property of the Laplace distribution of $\mathbf{Var}_{(K,\delta\mathcal{F}_k)}\left\{C'_i\right\} = 2\,b^2_{C_i}$ and **Lemma 3** is proved.

APPENDIX D
THE OPTIMAL FOURIER COEFFICIENTS NOISE SCALE PARAMETERS

Lemma 4: The optimal Fourier coefficients noise scale parameters for the constrained optimization problem of data utility $\mathcal{U}$ for a given privacy budget ε given in Eq. (10) are given by Eq. (15).

Proof: Starting from Eq. (10), we require that,

$$\left.\frac{\partial\mathcal{L}}{\partial b_{\mathcal{F}_k}}\right|_{b^*_{\mathcal{F}_k}} = 0,$$

from which we have,

$$\left(6\bar{\bar{I}} - \bar{\bar{Q}}\right)\overline{b^{*2}_{\mathcal{F}}} = \frac{1}{4\lambda^*}\overline{W'},$$

where $\bar{\bar{I}}$ is the unit matrix; $\overline{b^2_{\mathcal{F}}}$ is the vector of $\left\{b^2_{\mathcal{F}_k}|k = 0, 1, \ldots, K-1\right\}$; $\overline{W'}$ is the vector of $\left\{\mathcal{W}'_k|k = 0, 1, \cdots, K-1\right\}$; and the matrix $\bar{\bar{Q}}$ is given below,

$$\bar{\bar{Q}} = \begin{bmatrix}(1-p)^0 & (1-p)^1 & (1-p)^2 & (1-p)^3 & \cdots\\ \vdots & \vdots & \vdots & \vdots & \ldots\\ (1-p)^0 & (1-p)^1 & (1-p)^2 & (1-p)^3 & \cdots\end{bmatrix}.$$

First, let's express $\overline{b^{*2}_{\mathcal{F}}}$ as a function of λ^*,

$$\overline{b^{*2}_{\mathcal{F}}} = \frac{1}{\lambda^*}\frac{\left(6\bar{\bar{I}} - \bar{\bar{Q}}\right)^{-1}\overline{W'}}{4}. \tag{16}$$

Then, let's substitute Eq. (16) into Eq. (10) and obtain the analytical formula of λ^*,

At last, combing Eq. (16) and Eq. (17), as shown at the top of this page, we obtain the analytical formula for $b^*_{\mathcal{F}_k}$ given in Eq. (15), as shown at the top of this page, and thus **Lemma 4** is proved.

REFERENCES

[1] Y. Zheng, L. Z. Zhang, X. Xie, and W.-Y. Ma, "Mining interesting locations and travel sequences from GPS trajectories," in *Proc. WWW*, Madrid, Spain, 2009, pp. 791–800.

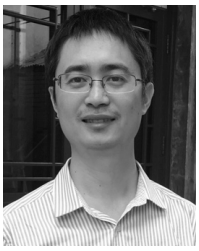
[2] J. Yuan, Y. Zheng, X. Xie, and G. Z. Sun, "Driving with knowledge from the physical world," in *Proc. SIGKDD*, Athens, Greece, 2011, pp. 316–324.

[3] H. Wang and Z. Xu, "CTS-DP: Publishing correlated time-series data via differential privacy," *Knowl.-Based Syst.*, vol. 122, pp. 167–179, Apr. 2017.
[4] T.-H. You, W.-C. Peng, and W.-C. Lee, "Protecting moving trajectories with dummies," in *Proc. MDM*, Beijing, China, May 2007, pp. 278–282.
[5] M. Terrovitis, G. Poulis, N. Mamoulis, and S. Skiadopoulos, "Local suppression and splitting techniques for privacy preserving publication of trajectories," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 7, pp. 1466–1479, Jul. 2017.
[6] R. Chen, B. C. M. Fung, N. Mohammed, B. C. Desai, and K. Wang, "Privacy-preserving trajectory data publishing by local suppression," *Inf. Sci.*, vol. 231, pp. 83–97, May 2013.
[7] L. Sweeney, "*k*-anonymity: A model for protecting privacy," *Int. J. Uncertainty Fuzziness Knowl. Based Syst.*, vol. 10, no. 5, pp. 557–570, 2002.
[8] S. R. Ganta, S. P. Kasiviswanathan, and A. Smith, "Composition attacks and auxiliary information in data privacy," in *Proc. SIGKDD*, Vancouver, BC, Canada, 2008, pp. 265–273.
[9] J. Hua, Y. Gao, and S. Zhong, "Differentially private publication of general time-serial trajectory data," in *Proc. INFOCOM*, Hong Kong, Apr./May 2015, pp. 549–557.
[10] C. Dwork, "Differential privacy," in *Proc. ICALP*, Venice, Italy, 2006, pp. 1–12.
[11] V. Rastogi and S. Nath, "Differentially private aggregation of distributed time-series with transformation and encryption," in *Proc. SIGMOD*, New York, NY, USA, 2010, pp. 735–746.
[12] K. Jiang, D. Shao, S. Bressan, T. Kister, and K. Tan, "Publishing trajectories with differential privacy guarantees," in *Proc. SSDBM*, Baltimore, MD, USA, 2013, Art. no. 12.
[13] R. Chen, B. C. M. Fung, B. C. Desai, and N. M. Sossou, "Differentially private transit data publication: A case study on the Montreal transportation system," in *Proc. KDD*, Beijing, China, 2012, pp. 213–221.
[14] R. Chen, G. Acs, and C. Castelluccia, "Differentially private sequential data publication via variable-length *n*-grams," in *Proc. CCS*, Raleigh, NC, USA, 2012, pp. 638–649.
[15] S. Wang and R. O. Sinnott, "Protecting personal trajectories of social media users through differential privacy," *Comput. Secur.*, vol. 67, pp. 142–163, Jun. 2017.
[16] D. Kifer and A. Machanavajjhala, "Pufferfish: A framework for mathematical privacy definitions," *ACM Trans. Database Syst.*, vol. 39, no. 1, Jan. 2014, Art. no. 3.
[17] S. Song, Y. Wang, and K. Chaudhuri, "Pufferfish privacy mechanisms for correlated data," in *Proc. SIGMOD*, Chicago, IL, USA, 2017, pp. 1291–1306.
[18] M. Guo, X. Jin, N. Pissinou, S. Zanlongo, B. Carbunar, and S. S. Iyengar, "In-network trajectory privacy preservation," *ACM CSUR*, vol. 48, no. 2, Oct. 2015, Art. no. 23.
[19] S. Gao, J. Ma, W. Shi, G. Zhan, and C. Sun, "TrPF: A Trajectory privacy-preserving framework for participatory sensing," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 6, pp. 874–887, Jun. 2013.
[20] D. Riboni and C. Bettini, "Differentially-private release of check-in data for venue recommendation," in *Proc. PerCom*, Budapest, Hungary, Mar. 2014, pp. 190–198.
[21] D. Quan, L. Yin, and Y. Guo, "Enhancing the trajectory privacy with Laplace mechanism," in *Proc. Trustcom*, Helsinki, Finland, Aug. 2015, pp. 1218–1223.
[22] G. Zhang, X. Liu, and Y. Yang, "Time-series pattern based effective noise generation for privacy protection on cloud," *IEEE Trans. Comput.*, vol. 64, no. 5, pp. 1456–1469, May 2015.
[23] Y. Cao and M. Yoshikawa, "Differentially private real-time data release over infinite trajectory streams," in *Proc. MDM*, Pittsburgh, PA, USA, Jun. 2015, pp. 68–73.
[24] M. Li, L. Zhu, Z. Zhang, and R. XU, "Achieving differential privacy of trajectory data publishing in participatory sensing," *Inf. Sci.*, vols. 400–401, pp. 1–13, Aug. 2017.
[25] V. Bindschaedler and R. Shokri, "Synthesizing plausible privacy-preserving location traces," in *Proc. SP*, San Jose, CA, USA, May 2016, pp. 546–563.
[26] X. He, G. Cormode, A. Machanavajjhala, C. M. Procopiuc, and D. Srivastava, "DPT: Differentially private trajectory synthesis using hierarchical reference systems," *Proc. VLDB Endowment*, vol. 8, no. 11, pp. 1154–1165, Jul. 2015.
[27] X. He, N. Raval, and A. Machanavajjhala, "A demonstration of VisDPT: Visual exploration of differentially private trajectories," *Proc. VLDB Endowment*, vol. 9, no. 13, pp. 1489–1492, Sep. 2016.
[28] L. Fan, L. Xiong, and V. S. Sunderam, "Differentially private multi-dimensional time series release for traffic monitoring," in *Proc. DBSec*, Newark, NJ, USA, 2013, pp. 33–48.
[29] L. Ou, Z. Qin, Y. Liu, H. Yin, Y. P. Hu, and H. Chen, "Multi-user location correlation protection with differential privacy," in *Proc. ICPADS*, Wuhan, China, Dec. 2017, pp. 422–429.
[30] A. A. Toda. (Nov. 2011). "Weak limit of the geometric sum of independent but not identically distributed random variables." [Online]. Available: https://arxiv.org/abs/1111.1786
[31] Y. Cao, M. Yoshikawa, Y. Xiao, and L. Xiong, "Quantifying differential privacy under temporal correlations," in *Proc. Int. Conf. Data Eng.*, Apr. 2017, pp. 821–832.
[32] Y. Chon, E. Talipov, H. Shin, and H. Cha, "Mobility prediction-based smartphone energy optimization for everyday location monitoring," in *Proc. SenSys*, Seattle, WA, USA, 2011, pp. 82–95.

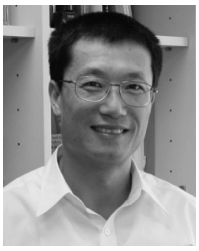
LU OU (S'15) received the B.S. degree in computer science from the Changsha University of Science and Technology and the M.S. degree in software engineering from Hunan University, China, in 2009 and 2012, respectively, where she is currently pursuing the Ph.D. degree with the College of Computer Science and Electronic Engineering. Her research focuses on security, privacy, and big data.

ZHENG QIN (M'18) received the Ph.D. degree in computer software and theory from Chongqing University, China, in 2001. From 2010 to 2011, he served as a Visiting Scholar with the Department of Computer Science, Michigan University. He is currently a Professor with the College of Computer Science and Electronic Engineering, Hunan University, where he also serves as the Vice Dean. He also serves as the Director of the Hunan Key Laboratory of Big Data Research and Application and the Vice Director of the Hunan Engineering Laboratory of Authentication and Data Security. His main interests are network and data security, privacy, data analytics and applications, machine learning, and applied cryptography. He is a member of the China Computer Federation.

SHAOLIN LIAO (SM'15) received the B.S. degree in materials science and engineering from Tsinghua University, Beijing, China, in 2000, and the Ph.D. degree in electrical engineering from the University of Wisconsin–Madison, Madison, USA, in 2008. He was a Post-Doctoral Fellow at the Department of Physics, City University of New York, from 2008 to 2010. He is currently a Research and Development Staff with the Argonne National Laboratory and an Adjunct Faculty with the Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, IL, USA. His interests span the multidisciplinary areas of privacy and machine learning of big data, simulation, algorithms, and modeling in signal processing, and novel methods in computational electromagnetics. He is an Associate Editor of the IEEE Access.



HUI YIN received the B.S. degree in computer science from Hunan Normal University, China, in 2002, the M.S. degree in computer software and theory from Central South University, China, in 2008, and the Ph.D. degree from the College of Information Science and Engineering, Hunan University, China, in 2018. He is currently an Assistant Professor with the College of Applied Mathematics and Computer Engineering, Changsha University, China. His main interests include information security, privacy protection, applied cryptography, and malware detection.



XIAOHUA JIA (F'13) received the B.Sc. and M.Eng. degrees from the University of Science and Technology of China in 1984 and 1987, respectively, and the D.Sc. degree in information science from The University of Tokyo in 1991. He is currently a Chair Professor with the Department of Computer Science, City University of Hong Kong. His research interests include cloud computing and distributed systems, data security and privacy, computer networks, and mobile computing. He is the General Chair of ACM MobiHoc 2008, a TPC Co-Chair of IEEE GLOBECOM 2010–Ad Hoc and Sensor Networking Symp, and an Area Chair of IEEE INFOCOM 2010 and 2015–2017. He is an Editor of the IEEE INTERNET OF THINGS, the IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS (2006–2009), *Wireless Networks*, the *Journal of World Wide Web*, the *Journal of Combinatorial Optimization*, and so on.