Quantifying Location Privacy for Navigation Services in Sustainable Vehicular Networks

Li, Meng; Chen, Yifei ; Kumar, Neeraj ; Lal, Chhagan; Conti, Mauro ; Alazab, Mamoun

DOI
10.1109/TGCN.2022.3144641

Publication date
2022

Document Version
Final published version

Published in
IEEE Transactions on Green Communications and Networking

Citation (APA)
Li, M., Chen, Y., Kumar, N., Lal, C., Conti, M., & Alazab, M. (2022). Quantifying Location Privacy for Navigation Services in Sustainable Vehicular Networks. *IEEE Transactions on Green Communications and Networking*, *6*(3), 1267-1275. https://doi.org/10.1109/TGCN.2022.3144641

Quantifying Location Privacy for Navigation Services in Sustainable Vehicular Networks

Meng Li, *Member, IEEE*, Yifei Chen, *Student Member, IEEE*, Neeraj Kumar, *Senior Member, IEEE*, Chhagan Lal, *Member, IEEE*, Mauro Conti, *Fellow, IEEE*, and Mamoun Alazab, *Senior Member, IEEE*

***Abstract*—Current connected and autonomous vehicles will contribute to various and green vehicular services. However, sharing personal data with untrustworthy Navigation Service Providers (NSPs) raises serious location concerns. To address this issue, many Location Privacy-Preserving Mechanisms (LPPMs) have been proposed. In addition, several quantification methods have been designed to help understand location privacy and illustrate how location privacy is leaked. However, their assessment is insufficient due to the incomplete assumptions about the adversary's model. In particular, users tend to request the same navigation routes from *home* to *workplace* and acquire traffic information along the route. An adversary can collect the coordinates of adjacent locations and infer the two true locations. In this paper, we provide a formal framework for the analysis of LPPMs in navigation services. Our framework captures extra information that is available to an adversary performing localization attacks. By formalizing the adversary's performance, we also propose and justify two new metrics to quantify location privacy in navigation services, namely *accuracy* and *visibility*. We assess the efficacy of two popular LPPMs for location privacy, i.e., differential privacy and *k*-anonymity. Experimental results demonstrate that the adversary can recover users' locations with a high probability.**

Manuscript received 8 June 2021; revised 24 July 2021 and 29 October 2021; accepted 17 January 2022. Date of publication 20 January 2022; date of current version 19 August 2022. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 62002094; in part by the Anhui Provincial Natural Science Foundation under Grant 2008085MF196; and in part by the EU LOCARD Project under Grant H2020-SU-SEC-2018-832735. *(Corresponding author: Mamoun Alazab.)*

Meng Li and Yifei Chen are with the Key Laboratory of Knowledge Engineering with Big Data, Ministry of Education, the School of Computer Science and Information Engineering, the Anhui Province Key Laboratory of Industry Safety and Emergency Technology, and the Intelligent Interconnected Systems Laboratory of Anhui Province, Hefei University of Technology, Hefei 230601, China (e-mail: mengli@hfut.edu.cn; yifeichen@mail.hfut.edu.cn).

Neeraj Kumar is with the Department of Computer Science Engineering, Thapar Institute of Engineering and Technology (Deemed to be University), Patiala 147004, India, also with the Department of Computer Science and Information Engineering, Asia University, Taichung City 413, Taiwan, also with the Department of Electrical and Computer Engineering, King Abdul Aziz University, Jeddah 80200, Saudi Arabia, and also with the Department of Computer Science, University of Petroleum and Energy Studies, Dehradun 248007, India (e-mail: neeraj.kumar@thapar.edu).

Chhagan Lal is with the Department of Intelligent Systems, Cyber Security Group, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: c.lal@tudelft.nl).

Mauro Conti is with the Department of Mathematics, University of Padua, 35131 Padua, Italy, and also with the Department of Intelligent Systems, Cyber Security Group, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: conti@math.unipd.it).

Mamoun Alazab is with the College of Engineering, IT and Environment, Charles Darwin University, Casuarina, NT 0810, Australia (e-mail: alazab.m@ieee.org).

Digital Object Identifier 10.1109/TGCN.2022.3144641

***Index Terms*—Vehicular networks, navigation services, location privacy, privacy quantification.**

I. Introduction

CURRENT vehicular networks have provided more efficient and green services [1]–[3] to vehicular users. With the development in communication technologies and hardware [4]–[7], connected and autonomous vehicles (CAV) will contribute significantly to the advancement of various vehicular services under green and sustainable economies.

Among the services, finding an optimal route from a given location to a destination is a common one for drivers. Due to the rapid development of sensing and communicating techniques, the increasing availability of users' locations has boosted the use of Location-Based Services (LBSs). In the LBS-based application market, navigation services have become favored in vehicular networks [8]–[12]. For example, Google Maps currently has more than 1 billion users worldwide [13].

In a typical navigation service, the user can be a driver maneuvering a vehicle equipped with an On-Board Unit (OBU) or a pedestrian holding a smartphone. The user sends a navigation request including her/his current location and a destination to a nearby Road-Side Units (RSUs) using Dedicated Short Range Communications protocol and Long Term Evolution (LTE)-based vehicle-to-everything technology (i.e., LTE-based V2X or 5G V2X). The RSU then forwards the navigation request to an NSP and returns a navigation route to the user. Such services offer better user experiences by allowing users to acquire optimal routes and driving guidance. These services also reduce traffic accidents and improve safety by enabling the vehicles to "see" the real-time road traffic.

While rendering convenience to users, sharing personal data (e.g., location) results in digital crumbs of their privacy [14]–[17]. This is primarily because the NSP is assumed to be an untrustworthy entity (or adversary) [18], [19] which leaks user information due to system malfunctioning or some malicious employee selling the information illegally. The consequences of location leakage are threefold. First, a user's visited locations are leaked. Sometimes, these visits are frequent which exposes more sensitive information, such as home. Second, a location is always correlated to an activity.

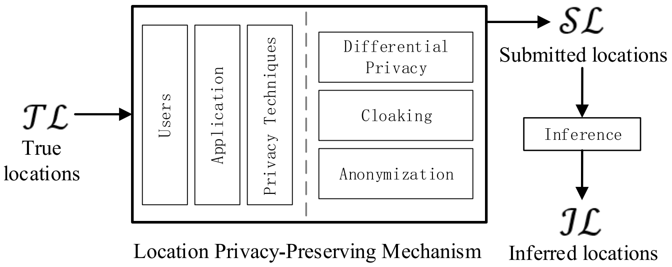


Fig. 1. Location privacy preserving mechanism in navigation services.

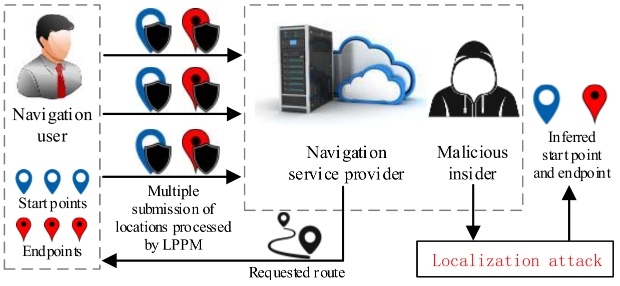

Navigation user
Start points
Endpoints
Multiple submission of locations processed by LPPM
Navigation service provider
Malicious insider
Inferred start point and endpoint
Requested route
Localization attack

Fig. 2. Localization attack against navigation services.

A visit to a hotel may reveal a secret meeting with someone. Third, combined with some background knowledge, an adversary will disclose the real identity of a user.

Note that a set of locations is more than a set of coordinates. Exposing this information can lead to three levels of location privacy leakage [20], [21]: 1) Primary level. The submitted locations directly reveal users' highly sensitive locations, e.g., home and workplace. 2) Deep level. Combined with the contextual information attached to the locations, the adversary can tell the users' habits and activities, e.g., eating weekly in an Italian restaurant and visiting a dentist five times in two months. 3) Interconnected level. If the adversary has acquired two sets of locations with the same spatiotemporal characters from two users, this will possibly disclose the relationship between these two users, e.g., two "strangers" exchanging business secrets in a coffee shop. Then, the location privacy leakage exposes the users to annoying advertisements, financial loss, and loss of time, as well as make them vulnerable to stalking and even criminal injury.

Several LPPMs have been proposed previously to protect location privacy, as depicted in Fig. 1, including the two popular differential privacy [22] and k-anonymity [23]. Differential privacy draws some noises from a distribution (e.g., a planar Laplace distribution [22]) and adds them to the locations before users upload their locations to the NSP, and k-anonymity refers to hiding the true location in a set of k locations realized by randomly selecting $k-1$ locations near the true locations. LPPMs function as a noisy channel that alters the location information communicated from users to the NSP [24]–[26]. They provide users a degree of control over the amount of location information shared with the NSP. In addition, governmental efforts have been made to promulgate privacy laws that address challenges raised by data sharing with commercial companies. For example, the European Union's General Data Protection Regulation (GDPR) went into effect on May 25, 2018. These laws enforce data autonomy by requiring the companies to increase transparency during data collection and strengthening user rights regarding their data.

The possibilities of attack surfaces have always been underestimated and defense mechanisms cannot adapt to new attacks efficiently. We have observed that many navigation users tend to request a route from home to workplace multiple times. For example, Alice uses Google Maps to request a route from her apartment to her place of employment on weekdays. No matter how the above-mentioned two mechanisms perturb or mask Alice's location in multiple requests, the submitted locations will form two general regions that remain stable as the number of requests increases. Integrated with some background knowledge, e.g., the layout of a community and commercial district, an untrustworthy NSP will eventually recover start points and endpoints, as shown in Fig. 2. We refer to this attack as *localization attack under multiple same requests*.

To quantify location privacy in LBSs, some contributions have been made in specific areas, e.g., building a unified framework for location privacy and defining components that affect location privacy [27], presenting a theoretical framework to model and quantify location privacy [28], developing a model to measure source-location information leakage for routing schemes in wireless sensor networks [29], inferring true identity of a user in a group of anonymous traces with some side information [30], and evaluating and configuring LPPMs [31]. These techniques have inspired us to quantify location privacy in navigation services. While we share common concerns, they did not provide a framework that can be used to quantify LPPM in navigation services under the new localization attack. Motivated by these issues, we propose to measure location privacy quantitatively when current LPPMs face a localization attack in navigation services. Taken as a whole, our work can be considered as an exploration of location privacy measurement in navigation services that can be used as a reference in designing privacy-preserving schemes. Our primary contributions are summarized as follows.

- We provide a generic model that formalizes the adversary's localization attacks against the homes and workplaces of navigation users.
- We propose two new metrics to quantify location privacy: 1) *accuracy*, i.e., the distance between the true location and inferred location and 2) *visibility* from the estimated location to the true location.
- We demonstrate the efficacy of two LPPMs, i.e., differential privacy and k-anonymity, when they are used to quantify location privacy in navigation services. We also present the success probability in recovering two frequently visited locations.

The remaining of this paper is organized as below. We give a formal description of our framework in Section II. We evaluate the efficacy of two LPPMs in navigation services, and provide a mitigation strategy in Section III. Some related work is discussed in Section IV. Lastly, we discuss some issues in Section V and conclude our work in Section VI.

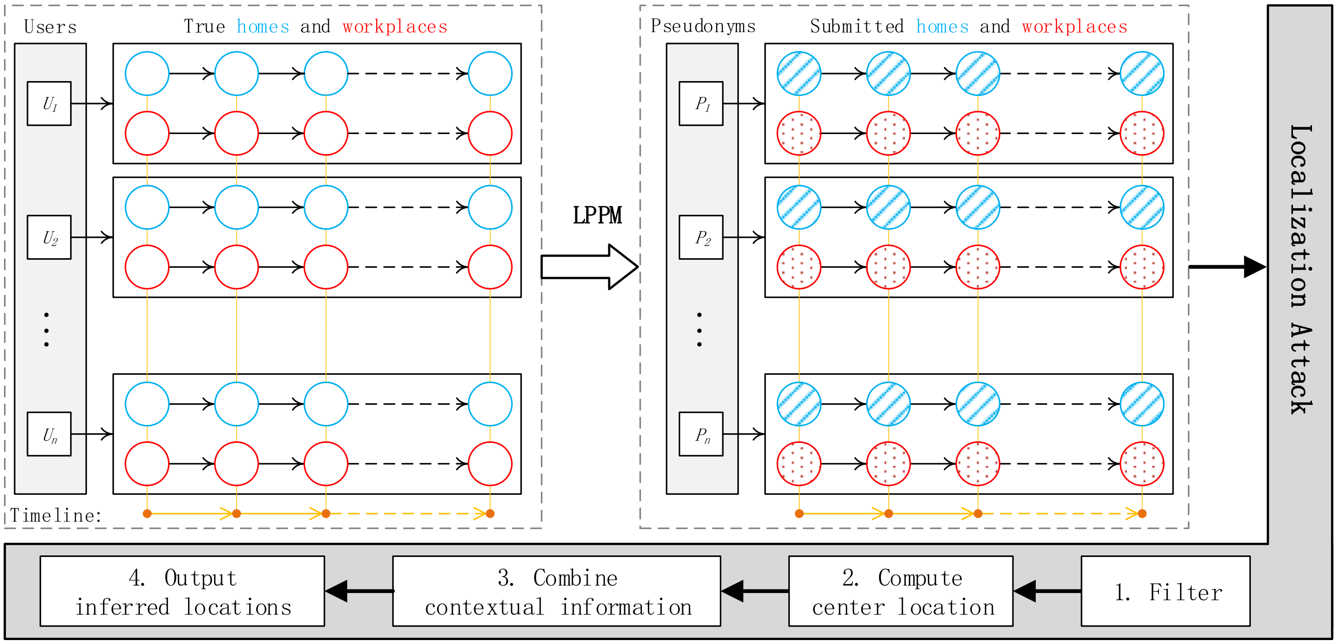

Users
True homes and workplaces
U_1
U_2
U_n
Timeline:
LPPM
Pseudonyms
Submitted homes and workplaces
P_1
P_2
P_n
Localization Attack
1. Filter
2. Compute center location
3. Combine contextual information
4. Output inferred locations

Fig. 3. Elements of the proposed location privacy framework.

II. Proposed Framework

We define a location-privacy framework as a tuple of elements: $(\mathcal{U}, \mathcal{TL}, \mathcal{T}, \text{LPPM}, \mathcal{P}, \mathcal{SL}, \mathcal{A}, \text{METRIC})$, where $\mathcal{U}$ is the set of navigation users, $\mathcal{TL}$ is the set of true locations, $\mathcal{T}$ is a set of time points at which the users submit a query, and LPPM represents the LPPM acting on true locations $tl \in \mathcal{TL}$ and producing submitted locations $sl \in \mathcal{SL}$. $\mathcal{P}$ is the set of pseudonyms with which users replace their real identities, and $\mathcal{SL}$ is the set of submitted locations. Here, adversary $\mathcal{A}$ is the NSP and an entity (eavesdropping on the communication channel) who implements the localization attack to infer *tl* after observing *sl* by relying on some background knowledge. The performance of $\mathcal{A}$ and its success probability in recovering true locations is characterized by an evaluation metric METRIC.

The framework is shown in Fig. 3, and the summary of the notations is presented in Table I. In the following subsections, we present and describe all elements of the proposed framework and discuss their interrelationship.

TABLE I
NOTATIONS

NSP	Navigation service provider
LPPM	Location Privacy-Preserving Mechanism
CAV	connected and autonomous vehicles
LBS	Location-based services
RSUs	Road-side unit
LTE	Long term evolution
GDPR	General data protection regulation
$\mathcal{U}$, $\mathcal{TL}$	Set of navigation users, set of true locations
$\mathcal{T}$, $\mathcal{P}$	Set of time points, set of pseudonyms
$\mathcal{SL}$, $\mathcal{A}$	Set of submitted locations, adversary
u, tl, t	User identity, true location, time point
p, sl, E	Pseudonym, submitted location, euclidean metric
f_1, f_2	Anonymization function, Perturbation function
M, r_{min}	Randomized mechanism, location radius
A_{min}, P	Privacy-aware cloaking region, location profile
PLI, H, W	Public location information, home, workplace
$\hat{H}$, $\hat{W}$	Inferred home, inferred workplace
HL, NL	Home locatio, noisy location
CL, N	Central location, number of noises
$d_{\widehat{CL}}$	Distance between HL and CL

A. Navigation Users

We denote $\mathcal{U} = \{u_1, u_2, \ldots, u_n\}$ a set of n navigation users who frequently query a route from point A to point B. Time is discrete and $\mathcal{T} = \{1, 2, \ldots, t\}$. Here, two sequential time points can be one minute, hour, or day apart.

B. LPPMs

The mechanism that modifies location to protect navigation users' location privacy is referred to as an LPPM, which processes navigation queries in two phases. In the anonymization phase, the identity of the querying user $u \in \mathcal{U}$ is replaced with a pseudonym $p \in \mathcal{P} = \{p_1, p_2, \ldots, p_n\}$ by using an anonymization function f_1. There are many ways to anonymize an identity, e.g., selecting a pseudonym randomly, computing a hash value of the concatenation of identity and timestamp, and using an anonymous credential. In the perturbation phase, the location of the querying user is perturbed to another location by using a perturbation function f_2.

Andrés *et al.* [22] proposed the Geo-Indistinguishability concept, where for two locations *x*, *y* and the Euclidean metric $\mathsf{E}(\cdot, \cdot)$, a randomized mechanism M satisfies $\mathsf{E}(\mathsf{M}(x), \mathsf{M}(y)) \leq \epsilon \mathsf{E}(x, y)$. Here, ϵ is referred to as the privacy budget which corresponds to the level of privacy.

Niu *et al.* [23] designed a virtual circle-based cloaking algorithm to satisfy *k*-anonymity. They first construct a virtual circle with a location radius r_{min} satisfying $\pi \cdot r_{min} \geq A_{min}$, where A_{min} is the privacy-aware cloaking region. The center c of the circle is selected randomly from the local map satisfying $r_{min} < \mathcal{E}(tl, c) \leq l/2$ where l is the length of the local map. The first perturbed location pl_1 is determined with an angle satisfying $\alpha_1 = \angle tlcpl_1 = 2\pi/k$. The other $k-1$ delusive locations can be selected around c clockwise in sequence.

LPPMs for navigation services can be implemented in different manners, i.e., centralized or distributed implementations. In the distributed implementation, the modification is performed by a trusted third party, i.e., a central anonymity server, as opposed to being performed by users on OBUs or smartphones in a distributed manner.

C. Adversary

When the navigation users submit their locations to the NSP, some curious entities can eavesdrop on the wireless communication channel to store these locations. A malicious employee at the NSP may leak locations to profit organizations for financial benefits. How we describe the new threat is very important to the proposed framework; therefore, we model adversary $\mathcal{A}$ prior to enforcing protection on location privacy. Here, $\mathcal{A}$ is portrayed by his background knowledge and localization attacks. $\mathcal{A}$ knows the anonymization function f_1 and perturbation function f_2, and he possesses some public location information, e.g., private residential community, apartment block, office building, and factory. Based on this information, $\mathcal{A}$ forms a location profile P_u for each user *u*.

When the attack type is different, the behavior of the system will shift accordingly. If not come up with a protection mechanism accordingly, privacy may be sacrificed. However, in this work, we concentrate on the location inferring attack.

D. Two Metrics to Quantify Location Privacy

The location profile P_u is the output of $\mathcal{A}$ attacking user *u*. Note that location profile P_u can range from general to specific. However, in the proposed framework, P_u contains two locations, i.e., an inferred home $\hat{H}$ and an inferred workplace $\hat{W}$. Adversary $\mathcal{A}$ cannot access infinite resources; thus, $\hat{H}$ and $\hat{W}$ are only estimates of two true locations *H* and *W*. The primary concern for a user is whether $\mathcal{A}$ finds the true locations. To illustrate the efficacy of how $\mathcal{A}$ attacks and quantify the location privacy, we present the following two metrics.

- *Accuracy:* We quantify the accuracy of each location in P_u by calculating the distance between $\hat{H}$ and *H*, i.e., $d_{\hat{H}} = \mathsf{E}(\hat{H}, H)$, and calculating the distance between $\hat{W}$ and *W*, i.e., $d_H = \mathsf{E}(\hat{W}, W)$. In other words, we estimate how the inferred locations deviate from the true locations and compute their distances. Here, a smaller distance indicates more successful localization.
- *Visibility:* We quantify the visibility of the true locations by calculating whether there is any barrier between $\hat{H}/\hat{W}$ and *H*/*W*. In other words, we estimate whether a person standing at the inferred locations can see the true locations. Here, higher visibility indicates more successful localization.

Note that accuracy by itself cannot cover the entirety of location privacy; thus, we propose to use visibility as a supplement. Given the same accuracy, e.g., $d_H = 5$ meters, if there is a high wall between *CL* and *HL*, then $\mathcal{A}$ cannot see *HL* or fully acquire the location privacy of *HL*, which makes the attack less effective.

III. Evaluation of LPPMs

In this section, we first introduce the database used in this study and present the experimental methodology. We then use two LPPMs (Section II-B), i.e., the differential privacy-based mechanism and *k*-anonymity based mechanism, and evaluate their efficacy on location privacy in navigation services.

A. Database

We use the GeoLife as the database [32], which was collected by the Microsoft Geolife project from April 2007 to October 2011. This dataset involves 182 users holding different GPS loggers and GPS-phones, and contains more than 17,000 trajectories with a total distance of 1,251,654 km. Each user has multiple files, each of which contains a GPS trajectory for a single day. A sequence of time-stamped points represents each trajectory, and it includes latitude, longitude, date, and time information. More than 90% of the trajectories were recorded densely, e.g., 1-5s apart. GeoLife tracks users' outdoor movements, including moving from home to workplace, which is suitable for our experiments.

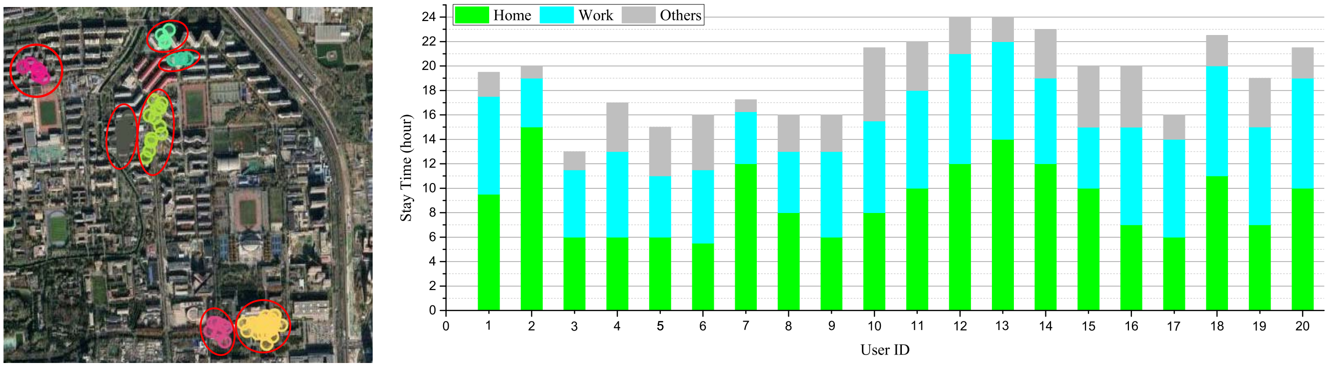
B. Experimental Methodology

First, we cluster the locations of each user by converting .plt files containing their separate GPS locations into .csv files, extracting the initial location item in the .csv files and employing a DBSCAN clustering algorithm. For example, as portrayed in Fig. 4(a), we found that user 000 had seven start location clusters marked in red circles after we clustered his/her start locations in 171 files.

Second, we manually review all location clusters for all 182 users and determined which stay points are homes and workplaces based on the duration of stay and public location information on the map. We set the stop time of each cluster to one hour to detect stay points that represent a Point-of-Interest (PoI). We then mark locations clusters on the Folium map using the *detection.stops()* function in the open-source library scikit-mobility (skmob). Then, we mark the location clusters in different colors based on the stay time to facilitate assessment. For example, if a user always stays in a location from 22:00 to 06:00, we consider this location to be that user's home. As illustrated in Fig. 4(b), we list the stay time and location clusters of 20 users for a typical day. The green pillar and blue pillar represent their home and workplace, which dominate their stay time. There are 61 and 68 users in the dataset who only show their home and workplace, respectively.

Third, we select a home location *HL* randomly from the abovementioned results as the home location. This location is set as the ground truth. Here, two LPPMs are used to generate noises.

Fourth, the noises are added to *HL* to produce noisy locations *NL*, and we compute the central location *CL* of the noisy locations are the inferred location and compute the distance $d_{\widehat{CL}}$ between *CL* and *HL*.



(a) 7 start location clusters of user 000

(b) Stay time of different location clusters for 20 users in one day

Fig. 4. Preprocessing of the dataset.

Finally, we determine any objects or barriers (e.g., fence, building, river) between *CL* and *HL*. To improve the accuracy of the inference attack, we combine some contextual information on the map and adjust *CL* to obtain a more reasonable output, i.e., the inferred location.

C. Differential Privacy-Based Mechanism

We quantify the location privacy of the differential privacy-based mechanism. The detailed experiment is as follows.

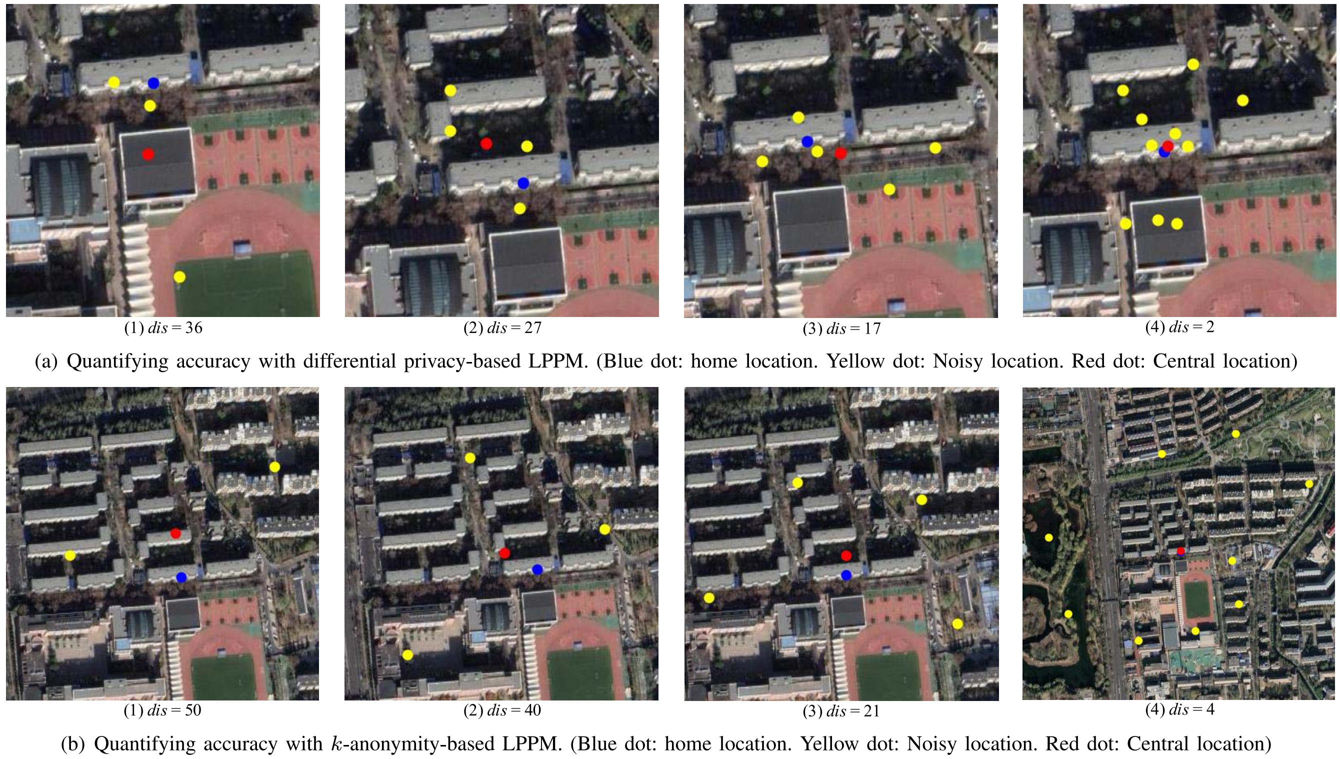
- We randomly select an *HL* for user 000 (40.0097, 116.3151).
- We draw N noises from the planar Laplace distribution whose probability density function is $\epsilon/2e^{-\epsilon|-x-\mu|}$ by using the *laplace()* function provided by the NumPy library. Here, μ is equal to 0, which means the drawn noises can be positive real numbers or negative real numbers. To achieve a meaningful perturbation and maintain the utility of the navigation service, we limit the noises to within [−10, 10], i.e., not more than 100 m away from the *HL*. It corresponds to the filter step in Fig. 3. Here, N noise values are 3, 4, 5, and 10.
- We add the N noises to the *HL* and obtain N noisy locations. Note that the coordinate system is *WGS*−84, and that meters have been converted to coordinate values on noisy values. We compute the *CL* of the N noisy locations.
- After obtaining the *CL*, we compute the distance $d_{\widehat{CL}}$ between *CL* and *HL* through using the *distance(CL, HL).m* method in GeoPy library. Here, the $d_{\widehat{CL}}$ indicates accuracy, i.e., smaller $d_{\widehat{CL}}$ values indicate higher accuracy.
- We identity *CL* and *HL* on the satellite map provided by the Folium library and Google Maps and check to see whether there is any object or barrier between *CL* and *HL* by manually observing the map. If there is no object or barrier, it means there is good visibility for $\mathcal{A}$ standing at *CL* toward *HL*.
- We combine the map information and adjust the *CL* to obtain a more reasonable inferred *HL*. For example, if the *CL* is 3 meters from an entrance gate of a private residential community, we consider the user's *HL* is this gate, and she/he lives in this community.

The experimental results are shown in Fig. 5(a), and we mark the *HL*, *NL*, and *CL* in blue, yellow, and red, respectively. We start from $N = 3$ because this scenario allows us to obtain a possible *CL*. Then, we observe how the $d_{\widehat{CL}}$ changes when $N = 4$ and $N = 5$. Note that the $d_{\widehat{CL}}$ decreases as N increases because the location privacy of the *CL* leaks more when there are more noisy locations around it. After N reaches 10, $d_{\widehat{CL}} = 2$. For visibility, we set $N \in [3, 10]$ and the resulting vector $\mathbf{v} = [1, 1, 1, 1, 0, 0, 0, 0]$ where 1 indicates there is an object or barrier, and 0 otherwise. Under these conditions, we can recover all *HL*s for all 182 users with a success probability of 100% when $N > 4$.

D. k-Anonymity-Based Mechanism

Here, we quantify the location privacy of the k-anonymity-based mechanism. The experimental details are as follows.

- Similar to the last subsection, we randomly select an *HL* for user 000 (40.0097, 116.3151).
- To maintain consistency with the previous experiment, we set r_{min} to 100. After the center c is selected, we construct a new *Cartesian Coordinate System* with c as the origin of the coordinate axis. Then, we use grid lines to split the coordinate system into cells where the cell size is 10∗10 (m^2).
- Given the values k and ρ, we implement a clockwise rotation algorithm to determine the next $k-1$ candidate locations. A geometric algorithm is also implemented to determine which cells are passed through each radial. For each radial, a passed through cell whose distance is no greater than $r * \rho$ from the candidate locations is randomly selected. The center of the selected cells will be noisy locations. Note, here k values are 3, 4, 5, and 10, and ρ is 0.5.
- After obtaining the $k-1$ noisy locations, we compute the center location *CL* of the $k-1$ noisy locations and the *HL* with the same method. For every location that needs to be shown on the map, its coordinates must be converted from the *Cartesian Coordinate System* to *WGS*−84.
- After obtaining *CL*, we also compute the distance between *HL* and *CL* by using *distance(CL, HL).m* method in the GeoPy library.



(a) Quantifying accuracy with differential privacy-based LPPM. (Blue dot: home location. Yellow dot: Noisy location. Red dot: Central location)

(b) Quantifying accuracy with k-anonymity-based LPPM. (Blue dot: home location. Yellow dot: Noisy location. Red dot: Central location)

Fig. 5. Quantifying location privacy.

- We show *CL* and *HL* on the satellite map and manually detect any object or barrier between *CL* and *HL*.
- Information around *CL* location is integrated to estimate a reasonable *HL*.

The experimental results are shown in Fig. 5(b). *HL*, *NL*, and *CL* are represented in blue, yellow, and red, respectively. As with the previous experiment, we start from $k = 3$ to obtain the *CL*. Then, we can determine whether $d_{\widehat{CL}}$ decreases as k increases. We find that when k increases to 10, $d_{\widehat{CL}}$ reaches 4. For the visibility, we also set $k \in [3, 10]$ and the resulting vector $\mathbf{v} = [1, 1, 1, 1, 0, 1, 0, 0]$. As the *HL* in this experiment is in a building, a reduction in the $d_{\widehat{CL}}$ does not necessarily change the visibility. However, when *HL* is in an open field, visibility will more likely be 0. In the end, we can recover 100% of the homes and workplaces for all the users after integrating some geographical information. This is because multiple releasing the same location with noises and background knowledge will disclose the location.

E. A Mitigation Strategy

To defend the localization attack, we could first turn the navigation model to a traffic congestion querying model, i.e., users do not have to input their home and workplace but only query the traffic status on certain roads. To protect location privacy, we could leverage secure searchable encryption. We will encrypt locations into secure indexes and users obtain the traffic by sending secure trapdoors to the NSP.

IV. Related Work

Shokri *et al.* [27] were the first to construct a unified framework for location privacy and define different components of location privacy. They modeled mobile networks (including users, time and space, and the spatiotemporal state of users), and the profile of users' activities after being processed by LPPMs. They identified three components related to protecting location privacy, i.e., users, applications, and privacy tools. Each component controls a certain amount of location information. The adversary is characterized by observing the output of an LPPM. It has three dimensions: means, actions, and goals. Next, they defined location privacy on macro and micro levels where the macro level referred to the user's privacy level throughout his trajectory and the micro level referred to the user's privacy on a small scale. By leveraging the proposed framework, they model location privacy in LBS and show its effectiveness.

Based on their previous work [27], Shokri *et al.* [28] presented a formal framework to quantify location privacy by formulating three location attacks, i.e., tracing attack, localization attack, and meeting disclosure attack, and proposed three quantification metrics, i.e., accuracy, certainty, and correctness. They pointed out that uncertainty and inaccuracy alone cannot measure location privacy. The core of the attack is whether the adversary reveals the correct answer or how close the adversary's estimation is to the answer. A distance, which is considered the correlation of the attack, can be computed by using the estimation and the answer. By adopting the existing statistical methods, they implemented attacks to measure user location privacy and evaluate the efficacy of entropy and k-anonymity.

Li *et al.* [29] presented a model and three criteria to quantify source-location privacy in existing routing-based schemes for wireless sensor networks. For a routing traceback attack and to reduce a source node space attack, they define two criteria to quantify the leakage of source-location information, i.e., source-location disclosure index and source-location space index. An additional criterion is the normalized source-location space index. These metrics can also be used in many energy-constrained applications.

Ma *et al.* [30] focused on how an adversary with some side information can infer an extended view of the locations of a user in an anonymous trajectory. Their experimental results quantify the loss of user privacy as a function of several parameters, including node mobility, inference strategies, and noises in the trajectory or side information. To some extent, the side information implies the location of a user at a certain time; however, the implied information may not be accurate. In reality, the side information could be acquired by chance or engineered encounters. Different strategies are used to identify the users' trajectories, such as the maximum likelihood estimator based approach and a minimum square approach. To measure the efficacy of strategies, three metrics are defined, i.e., fraction of correct conclusions, fraction of incorrect conclusions, and fraction of undecided conclusions.

Primault *et al.* [31] proposed a framework ALP to support evaluation and configuration of LPPMs. It forms a generic model to specify privacy and utility goals that an LPPM should satisfy. Instead of setting static configuration parameters, it designs an optimizer to adjust the parameters to meet the privacy and utility goals. However, the number of available metrics and the definition of the objectives are limited.

Besides, there are cryptography-based approaches in protecting locations in vehicular networks and social networks. Li and Jung [33] tackled the challenge raised from guaranteeing location privacy and utility at the same time. They proposed a new fine-grained privacy-preserving location query scheme based on attribute-based encryption and functional encryption to achieve different levels of location query for mobile platforms. Puttaswamy *et al.* [34] utilized secure user-specific and distance-preserving coordinate transformations to all locations shared with the cloud server for geosocial applications. A friend shares a user's secrets so that they can use the same transformation. It enables all location queries to be processed correctly by the cloud server while the server cannot know or infer the actual locations. Yu *et al.* [35] proposed a privacy-preserving protocol to exploit the sparse meeting opportunities for pseudonym changing in vehicular social networks. They leverage group signatures to build pseudonym-changing regions where vehicles exchange their pseudonyms. It enlarges the uncertainty of pseudonym mixture for tracking adversaries.

V. Discussions

A. Business Model

Since current commercial corporations are more concentrated on utility and efficiency, enforcing location privacy-preserving mechanisms will incur computational burdens. It is necessary to convince them to use such mechanisms. There are some options for us to choose. First, we can use lightweight cryptography [36], [37] to reduce computational costs. Second, local differential privacy is also applied by Google Chrome to protect users it barely affect the efficiency while keep a reasonable utility. Furthermore, with the improvement of device capabilities, the extra burdens from location privacy-preserving mechanisms will not be a significant issue.

B. Future Research Directions

The future research directions include several aspects. First, semantic privacy in location privacy is to be explored and measured. Second, it is necessary to locate other possible sources of location privacy leakage, such as users' misoperations and cross-references with other social platforms. Third, user-defined privacy should be integrated into the protection mechanism to meet different user requirements. Next, privacy computing that combines several existing privacy-enhancing techniques is a powerful tool to further enhance location privacy. Last, a united metric for quantifying privacy and pertinent standardization are in need.

VI. Conclusion

Location information is of the utmost importance in many LBSs because location information correlates to both system utility and user privacy. This work targets the navigation service in vehicular networks and quantifies users location privacy under LPPMs. First, we provide a formal description of our framework and propose two new metrics, i.e., accuracy and visibility, to quantify location privacy. Then, we evaluate the efficacy of two LPPMs in navigation services, i.e., differential privacy and *k*-anonymity. From the experimental results by using a real-world dataset, we show that we can infer users' home and workplace if they frequently query the route between the two locations.

Evaluating the efficiency of an LPPM and quantifying location privacy is no easy task. This study explores location privacy measurement customized for navigation services. We suggest that special attention be paid to privacy quantification in LBSs when designing LPPMs for such services.

Acknowledgment

This work was carried out during the tenure of an ERCIM 'Alain Bensoussan' Fellowship Programme granted to Dr. Meng Li.

References

[1] K. Dev, R. K. Poluru, R. L. Kumar, P. K. R. Maddikunta, and S. A. Khowaja, "Optimal radius for enhanced lifetime in IoT using hybridization of rider and grey wolf optimization," *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 2, pp. 635–644, Jun. 2021.

[2] A. A. Shah, N. A. Bhatti, K. Dev, and B. S. Chowdhry, "MUHAFIZ: IoT-based track recording vehicle for the damage analysis of the railway track," *IEEE Internet Things J.*, vol. 8, no. 11, pp. 9397–9406, Jun. 2021.

[3] C. D. Alwis *et al.*, "Survey on 6G frontiers: Trends, applications, requirements, technologies and future research," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 836–886, 2021.

[4] M. Alazab, K. Lakshmanna, G. T. Reddy, Q.-V. Pham, and P. K. R. Maddikunta, "Multi-objective cluster head selection using fitness averaged rider optimization algorithm for IoT networks in smart cities," *Sustain. Energy Technol. Assess.*, vol. 43, Feb. 2021, Art. no. 100973.

[5] A. Rehman, S. U. Rehman, M. Khan, M. Alazab, and G. T. Reddy, "CANintelliIDS: Detecting in-vehicle intrusion attacks on a controller area network using CNN and attention-based GRU," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 2, pp. 1456–1466, Apr.–Jun. 2021.

[6] M. Li, L. Zhu, and X. Lin, "Efficient and privacy-preserving carpooling using blockchain-assisted vehicular fog computing," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4573–4584, Jun. 2019, doi: 10.1109/JIOT.2018.2868076.

[7] M. Li, D. Hu, C. Lal, M. Conti, and Z. Zhang, "Blockchain-enabled secure energy trading with verifiable fairness in Industrial Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 16, no. 10, pp. 6564–6574, Oct. 2020, doi: 10.1109/TII.2020.2974537.

[8] 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, doi: 10.1016/j.ins.2017.03.015.

[9] L. Zhu, M. Li, Z. Zhang, and Z. Qin, "ASAP: An anonymous smart-parking and payment scheme in vehicular networks," *IEEE Trans. Dependable Secure Comput.*, vol. 17, no. 4, pp. 703–715, Jul./Aug. 2020, doi: 10.1109/TDSC.2018.2850780.

[10] M. Li, L. Zhu, and X. Lin, "Privacy-preserving traffic monitoring with false report filtering via fog-assisted vehicular crowdsensing," *IEEE Trans. Services Comput.*, vol. 14, no. 6, pp. 1902–1913, Nov./Dec. 2021, doi: 10.1109/TSC.2019.2903060.

[11] M. Li, Y. Chen, S. Zheng, D. Hu, C. Lal, and M. Conti, "Privacy-preserving navigation supporting similar queries in vehicular networks," *IEEE Trans. Dependable Secure Comput.*, early access, Aug. 18, 2020, doi: 10.1109/TDSC.2020.3017534.

[12] M. Li, Y. Chen, C. Lal, M. Conti, M. Alazab, and D. Hu, "Eunomia: Anonymous and secure vehicular digital forensics based on blockchain," *IEEE Trans. Dependable Secure Comput.*, early access, Nov. 25, 2021, doi: 10.1109/TDSC.2021.3130583.

[13] K. Wiggers. "Google Maps Turns 15 With a New Icon and Detailed Transit Data." Feb. 2020. [Online]. Available: https://venturebeat.com/2020/02/06/google-celebrates-maps-15th-birthday-with-a-new-icon-and-detailed-transit-data accessed Jul. 17, 2020.

[14] M. Li, L. Zhu, and X. Lin, "CoRide: A privacy-preserving collaborative-ride hailing service using blockchain-assisted vehicular fog computing," in *Proc. ACM 15th EAI Int. Conf. Security Privacy Commun. Netw. (SecureComm)*, Oct. 2019, pp. 408–422.

[15] L. Zhu, M. Li, and Z. Zhang, "Secure fog-assisted crowdsensing with collusion resistance: From data reporting to data requesting," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5473–5484, Jun. 2019, doi: 10.1109/JIOT.2019.2902459.

[16] M. Tariq, F. Naeem, M. Ali, and H. V. Poor, "Vulnerability assessment of 6G enabled smart grid cyber-physical systems," *IEEE Internet Things J.*, vol. 8, no. 7, pp. 5468–5475, Apr. 2021.

[17] M. Kamal, G. Srivastava, and M. Tariq, "Blockchain based lightweight and secured V2V communication in Internet of Vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 3997–4004, Jul. 2021.

[18] M. Li, F. Wu, G. Chen, L. Zhu, and Z. Zhang, "How to protect query and report privacy without sacrificing service quality in participatory sensing," in *Proc. 34th IEEE Int. Perform. Comput. Commun. Conf. (IPCCC)*, Dec. 2015, pp. 1–7.

[19] M. Li, L. Zhu, Z. Zhang, and R. Xu, "Differentially private publication scheme for trajectory data," in *Proc. 1st IEEE Int. Conf. Data Sci. Cyberspace (DSC)*, Jun. 2016, pp. 596–601.

[20] M. Li, L. Zhu, Z. Zhang, C. Lal, M. Conti, and F. Martinelli, "Privacy for 5G-supported vehicular networks," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 1935–1956, 2021.

[21] M. Li, L. Zhu, Z. Zhang, C. Lal, M. Conti, and M. Alazab, "Anonymous and verifiable reputation system for E-commerce platforms based on blockchain," *IEEE Trans. Netw. Service Manage.*, vol. 18, no. 4, pp. 4434–4449, Dec. 2021.

[22] M. E. Andrés, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, "Geo-indistinguishability: Differential privacy for location-based systems," in *Proc. 20th ACM Conf. Comput. Commun. Security (CCS)*, Nov. 2013, pp. 901–914.

[23] B. Niu, Z. Zhang, X. Li, and H. Li, "Privacy-area aware dummy generation algorithms for location-based services," in *Proc. Int. Conf. Commun.*, Jun. 2014, pp. 957–962.

[24] Y. Chen, M. Li, S. Zheng, D. Hu, C. Lai, and M. Conti, "One-time, oblivious, and unlinkable query processing over encrypted data on cloud," in *Proc. 22nd Int. Conf. Inf. Commun. Security (ICICS)*, Aug. 2020, pp. 350–365.

[25] M. Li, C. Lal, M. Conti, and D. Hu, "LEChain: A blockchain-based lawful evidence management scheme for digital forensics," *Future Gener. Comput. Syst.*, vol. 115, pp. 406–420, Feb. 2021, doi: 10.1016/j.future.2020.09.038.

[26] M. Li, J. Gao, Y. Chen, J. Zhao, and M. Alazab, "Privacy-preserving ride-hailing with verifiable order-linking in vehicular networks," in *Proc. 19th Int. Conf. Trust Security Privacy Comput. Commun. (TrustCom)*, Dec. 2020, pp. 599–606.

[27] R. Shokri, J. Freudiger, J.-P. Hubaux, "A unified framework for location privacy," in *Proc. 3rd Hot Topics Privacy Enhanc. Technol. (HotPETs)*, Sep. 2010, pp. 1–21.

[28] R. Shokri, G. Theodorakopoulos, J.-Y. Le Boudec, and J.-P. Hubaux, "Quantifying location privacy," in *Proc. 32rd IEEE Symp. Security Privacy (S&P)*, May 2011, pp. 247–262.

[29] Y. Li, J. Ren, and J. Wu, "Quantitative measurement and design of source-location privacy schemes for wireless sensor networks," *IEEE Trans. Parallel Distrib. Comput.*, vol. 23, no. 7, pp. 1302–1311, Jul. 2012.

[30] C. Y. T. Ma, D. K. Y. Yau, N. K. Yip, and N. S. V. Rao, "Privacy vulnerability of published anonymous mobility traces," *IEEE/ACM Trans. Netw.*, vol. 21, no. 3, pp. 720–733, Jun. 2013.

[31] V. Primault, A. Boutet, S. Ben Mokhtar, and L. Brunie, "Adaptive location privacy with ALP," in *Proc. 35th IEEE Symp. Reliable Distrib. Syst. (SRDS)*, Sep. 2016, pp. 269–278.

[32] Y. Zheng, H. Fu, X. Xie, W.-Y. Ma, and Q. Li, "Geolife GPS Trajectory Dataset—User Guide." 2011. [Online]. Available: https://www.microsoft.com/en-us/research/publication/geolife-gps-trajectory-dataset-user-guide accessed May 16, 2021.

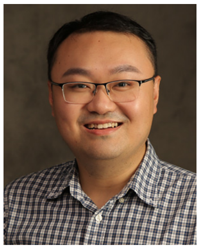
[33] X.-Y. Li and T. Jung, "Search me if you can: Privacy-preserving location query service," in *Proc. 32nd Annu. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, 2013, pp. 2760–2768.

[34] K. P. N. Puttaswamy *et al.*, "Preserving location privacy in geosocial applications," *IEEE Trans. Mobile Comput.*, vol. 13, no. 1, pp. 159–173, Jan. 2014.

[35] R. Yu, J. Kang, X. Huang, S. Xie, Y. Zhang, and S. Gjessing, "MixGroup: Accumulative pseudonym exchanging for location privacy enhancement in vehicular social networks," *IEEE Trans. Dependable Secure Comput.*, vol. 13, no. 1, pp. 93–105, Jan./Feb. 2016.

[36] Y. Zhang, C. C. Tan, F. Xu, H. Han, and Q. Li, "VProof: Lightweight privacy-preserving vehicle location proofs," *IEEE Trans. Veh. Technol.*, vol. 64, no. 1, pp. 378–385, Jan. 2015.

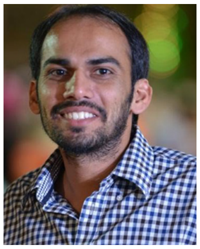
[37] Y. Naito, Y. Sasaki, and T. Sugawara, "Lightweight authenticated encryption mode suitable for threshold implementation," in *Proc. 39th Annu. Int. Conf. Theory Appl. Cryptogr. Techn. (EUROCRYPT)*, May 2020, pp. 705–735.



Meng Li (Member, IEEE) received the B.E. degree in the information security from the Hefei University of Technology in 2010, the M.S. and Ph.D. degrees in computer science and technology from the Beijing Institute of Technology in 2013 and 2019, respectively. He is currently an Associate Researcher and the Dean Assistant with the School of Computer Science and Information Engineering, Hefei University of Technology. He is also a Postdoctoral Fellow with the Department of Mathematics and HIT Center, University of Padua, Italy, where he is with the Security and PRIvacy Through Zeal Research Group led by Prof. M. Conti. He was sponsored by ERCIM "Alain Bensoussan" Fellowship Programme in October 2019 to conduct Postdoctoral Research with CNR, Italy. He was sponsored by China Scholarship Council to study in the Broadband Communications Research Lab, University of Waterloo and Wilfrid Laurier University from September 2017 to August 2018. His research interests include security, privacy, vehicular networks, applied cryptography, and blockchain. In this area, he has published more than 40 papers in international peer-reviewed transactions, journals, magazines, and conferences, including IEEE Transactions on Dependable and Secure Computing, IEEE Transactions on Services Computing, IEEE Transactions on Network and Service Management, IEEE Transactions on Industrial Informatics, IEEE Transactions on Network Science and Engineering, IEEE Internet of Things Journal, *Journal of Information Science*, *IEEE Communications Magazine*, IEEE Wireless Communications, MobiCom, ICICS, SecureComm, TrustCom, and IPCCC.

Yifei Chen (Student Member, IEEE) received the B.E. degree from the Hefei University of Technology, Hefei, China, in 2019, where he is currently pursuing the M.S. degree with the School of Computer Science and Information Engineering. His research interests include applied cryptography, security and privacy, vehicular networks, android development, and blockchain.

Neeraj Kumar (Senior Member, IEEE) received the Ph.D. degree in CSE from Shri Mata Vaishno Devi University, Katra, India. He was a Postdoctoral Research Fellow with Coventry University, Coventry, U.K. He is working as an Associate Professor with the Department of Computer Science and Engineering, Thapar Institute of Engineering and Technology (Deemed to be University), Patiala, India. He has published more than 450 technical research papers in leading journals and conferences from IEEE, Elsevier, Springer, and John Wiley. He is on the editorial board of *ACM Computing Surveys*, IEEE TRANSACTIONS ON SUSTAINABLE COMPUTING, *IEEE Network Magazine*, IEEE COMMUNICATIONS MAGAZINE, *Journal of Network and Computer Applications* (Elsevier), and *Computer Communications* (Elsevier).



Chhagan Lal (Member, IEEE) received the Ph.D. degree in computer science and engineering from the Malaviya National Institute of Technology, India, in 2014. He is a Senior Researcher of CyberSecurity with the Faculty of Electrical Engineering, Mathematics and Computer Science, Department of Intelligent Systems, Delft University of Technology, The Netherlands. Earlier, he was a Research Scientist with Simula Research Labs, Oslo, Norway. Before joining Simula, he was a Postdoctoral Fellow with the Department of Mathematics, University of Padova, Italy, where he is an active member of the Security and PRIvacy Through Zeal Research Group, which is led by Prof. M. Conti. His current research areas include applications of blockchain technologies, security in software-defined networking, and Internet of Things networks. During his Ph.D., he has been awarded with the Canadian Commonwealth Scholarship under the Canadian Commonwealth Scholarship Program to work in University of Saskatchewan, Saskatoon, SK, Canada.

Mauro Conti (Fellow, IEEE) received the Ph.D. degree from the Sapienza University of Rome, Italy, in 2009. He is a Full Professor with the University of Padua, Italy. He is also affiliated with the TU Delft and University of Washington, Seattle. After his Ph.D., he was a Postdoctoral Researcher with Vrije Universiteit Amsterdam, The Netherlands. In 2011, he joined as an Assistant Professor with the University of Padua, where he became an Associate Professor in 2015, and a Full Professor in 2018. He has been a Visiting Researcher with the GMU, UCLA, UCI, TU Darmstadt, UF, and FIU. His research is also funded by companies, including Cisco, Intel, and Huawei. His main research interest is in the area of Security and Privacy. In this area, he published more than 400 papers in topmost international peer-reviewed journals and conferences. He has been awarded with a Marie Curie Fellowship in 2012 by the European Commission, and with a Fellowship by the German DAAD in 2013. He is the Editor-in-Chief for IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, the Area Editor-in-Chief for IEEE COMMUNICATIONS SURVEYS AND TUTORIALS, and has been an Associate Editor for several journals, including IEEE COMMUNICATIONS SURVEYS AND TUTORIALS, IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, and IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT. He was a Program Chair for TRUST 2015, ICISS 2016, WiSec 2017, ACNS 2020, and CANS 2021, and a General Chair for SecureComm 2012, SACMAT 2013, NSS 2021, and ACNS 2022. He is a Senior Member of the ACM and a Fellow of the Young Academy of Europe.

Mamoun Alazab (Senior Member, IEEE) received the Ph.D. degree in computer science from the School of Science, Information Technology and Engineering, Federation University of Australia. He is an Associate Professor with the College of Engineering, IT and Environment, Charles Darwin University, Australia. He is a Cybersecurity Researcher and a Practitioner with industry and academic experience. He works closely with government and industry on many projects, including the Northern Territory, Department of Information and Corporate Services, IBM, Trend Micro, the Australian Federal Police, the Australian Communications and Media Authority, Westpac, United Nations Office on Drugs and Crime, and the Attorney Generals Department. He delivered many invited and keynote speeches, 24 events in 2019 alone. His research is multidisciplinary that focuses on cybersecurity and digital forensics of computer systems with a focus on cybercrime detection and prevention. He has more than 150 research papers. He convened and chaired more than 50 conferences and workshops. He is the Founding Chair of the IEEE NT Subsection.