A Privacy-Preserving Querying Mechanism with High Utility for Electric Vehicles

Ugur Ilker Atmaca, Sayan Biswas, Carsten Maple, Catuscia Palamidessi

IEEE Open Journal of Vehicular Technology

Received 11 October 2023; accepted 21 January 2024. Date of publication 30 January 2024;
date of current version 20 February 2024. The review of this article was coordinated by Editor Miao Pan.

Digital Object Identifier 10.1109/OJVT.2024.3360302

A Privacy-Preserving Querying Mechanism with High Utility for Electric Vehicles

UGUR ILKER ATMACA [1,2], SAYAN BISWAS [3,4,5], CARSTEN MAPLE [1,2], AND CATUSCIA PALAMIDESSI [3,4] (Member, IEEE)

[1]WMG, The University of Warwick, CV4 7AL Coventry, U.K.
[2]The Alan Turing Institute, NW1 2DB London, U.K.
[3]INRIA Saclay, 91128 Palaiseau, France
[4]LIX, École Polytechnique, 91120 Palaiseau, France
[5]EPFL, 1015 Lausanne, Switzerland

CORRESPONDING AUTHORS: UGUR ILKER ATMACA; SAYAN BISWAS (e-mail: ugur-ilker.atmaca@warwick.ac.uk; sayan.biswas@epfl.ch)

This work was supported in part by the Academic Centre of Excellence in Cyber Security Research - University of Warwick under Grant EP/R007195/1, in part by The Alan Turing Institute under Grant EP/N510129/1, in part by the PETRAS National Centre of Excellence for IoT Systems Cybersecurity under Grant EP/S035362/1, and in part by Autotrust under Grant EP/R029563/1. *(Ugur Ilker Atmaca and Sayan Biswas contributed equally to this work.)*

ABSTRACT Electric vehicles (EVs) are becoming more popular due to environmental consciousness. The limited availability of charging stations (CSs), compared to the number of EVs on the road, has led to increased range anxiety and a higher frequency of CS queries during trips. Simultaneously, personal data use for analytics is growing at an unprecedented rate, raising concerns for privacy. One standard for formalising location privacy is geo-indistinguishability as a generalisation of local differential privacy. However, the noise must be tuned properly, considering the implications of potential utility losses. In this paper, we introduce the notion of approximate geo-indistinguishability (AGeoI), which allows EVs to obfuscate their query locations while remaining within their area of interest. It is vital because journeys are often sensitive to a sharp drop in quality of service (QoS). Our method applies AGeoI with dummy data generation to provide two-fold privacy protection for EVs while preserving a high QoS. Analytical insights and experiments demonstrate that the majority of EVs get "privacy-for-free" and that the utility loss caused by the gain in privacy guarantees is minuscule. In addition to providing high QoS, the iterative Bayesian update allows for a private and precise CS occupancy forecast, which is crucial for unforeseen traffic congestion and efficient route planning.

INDEX TERMS Charging station, electric vehicle, geo-indistinguishability, location privacy, privacy-utility trade-off.

I. INTRODUCTION

Air pollution is one of the immediate issues that the world is experiencing [1], [2], [3]. In the United Kingdom in 2019, 27% of all greenhouse gas emissions came from transportation, as the largest emitting sector [4], [5], [6]. Hence, the transportation industry and academic communities are increasingly interested in developing alternative energy vehicles to reduce emissions. Automobile manufacturers are introducing a new generation of electric vehicles (EVs) that often employ connected and automated driving functions [7].

EVs are regarded as one of the most promising means of reducing emissions and reliance on fossil fuels. Along with environmental benefits, EVs provide superior energy efficiency to conventional vehicles [8]. As the cost of batteries continues to decrease, the large-scale adoption of EVs is becoming more viable [9]. Despite the advantages and competitive cost, many customers remain concerned about running out of battery power before reaching their destination or waiting for their EVs to charge. The primary obstacles to EV adoption are the availability of chargers and the range that can be travelled on a single charge, often referred to as *range anxiety* in the literature [10].

There has been some recent focus on forecasting how busy the charging stations (CS) are in certain areas to ensure that the EVs can plan their journeys conveniently [11], [12]. However, the existing research in this direction, primarily founded upon machine learning based methods, does not address the privacy concerns involved in such predictive techniques and

does not consider situations where there may arise unprecedented traffic congestion (e.g. due to a one-off concert or an event). One of the most successful approaches for protecting the privacy of personal data while analysing and exploiting the utility of data *differential privacy* (DP) [13], [14], which mathematically guarantees that the query output does not change significantly regardless of whether a specific personal record is in the dataset or not. Our proposed method, in addition to allowing the EVs to have formal privacy guarantees on their queries to locate the nearest CS, enables the users to estimate the live occupancy of the CS efficiently allowing convenient journey planning.

However, the classical central DP requires a trusted curator who is responsible for adding noise to the data before publishing or performing analytics on it. A major drawback of such a central model is that it is vulnerable to security breaches because the entire original data is stored in a central server. Moreover, there is the risk of having an adversarial curator. To circumvent the need for such a central dependency, a local model of DP, also called *local differential privacy (LDP)* [15], has been getting a lot of attention lately. In this model, users apply the LDP mechanism directly to their data and send the locally changed data to the server.

LDP is particularly suitable for situations where users need to communicate their personal data in exchange for some service. One such scenario is the use of location-based services (LBS), where a user typically reports her location in exchange for information like the shortest path to a destination, points of interest in the surroundings, traffic information, friends nearby, etc. One of the recently popularised standards in location privacy is *geo-indistinguishability* (GeoI) [16], which optimises the quality of service (QoS) for users while preserving a generalised notion of LDP on their location data. The obfuscation mechanism of GeoI depends on the distance between the original location of a user and a potential noisy location that they report [17], [18]. GeoI can be implemented directly on the user's device (tablet, smartphone, etc.). The fact that the users can control their explicit privacy-protection level for various LBS makes it very appealing. However, a drawback of injecting noise locally to the datum is that it deteriorates the QoS due to the lack of accuracy of the data.

On the other hand, future vehicles are becoming more sophisticated in their sensory, onboard computation, and communication capacities. Furthermore, the emergence of Mobile Edge Computing (MEC) also changes the Intelligent Transportation Systems (ITS) by providing a platform to assist computationally heavy tasks by offloading the computation to the Edge cloud [19]. This architecture often employs three tiers, with the vehicle on the first, MEC on the second, and standard cloud services on the third [20]. Fig. 1 shows the system architecture for the location privacy framework proposed in this paper.

ITS provides a platform containing distributed and resource-constrained systems to support real-time vehicular functions where these functions' efficacy relies on the data

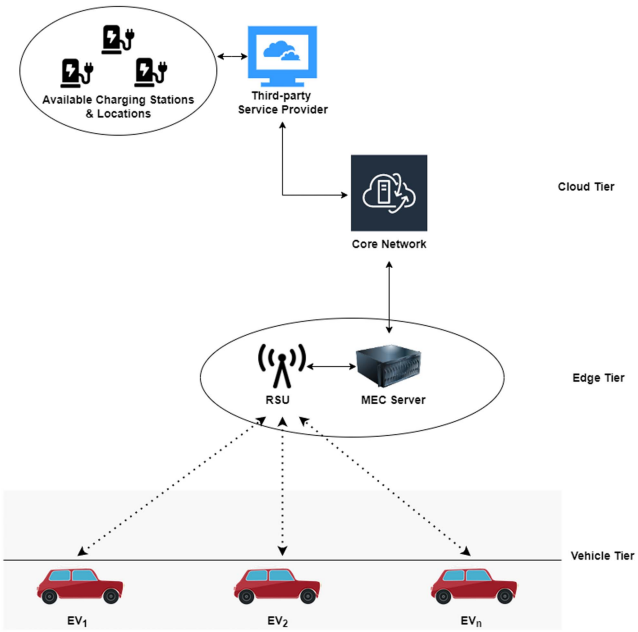


FIGURE 1. System architecture (EV: Electric Vehicle, RSU: Roadside Unit, MEC: Mobile-Edge Computing Unit).

shared across entities. However, the risk of privacy disclosure and tracking increases due to data sharing [21]. Privacy-preserving schemes are developed using established techniques such as group signature, anonymity, and pseudonymity [22], [23]. However, it is possible to identify privatised data with adequate background information. Hence, DP approaches have emerged as the gold standard of data privacy because they provide a formal privacy guarantee independent of a threat actor's background knowledge and computing capability [24].

GeoI is the state-of-the-art method for location privacy-preserving with LDP. It can preserve one's location privacy among a set of locations with similar probability distributions without requiring a trusted third-party. It provides rigorous privacy for location-based query processing and location data collection by modelling the location domain based on the Euclidean plane. However, vehicles are located on the road network under normal circumstances. For vehicular location queries, GeoI mechanism may result in publishing unrealistic privatised locations such as houses, parks, or lakes. Thus, there is a need for an adapted model of GeoI for vehicular application. This paper proposes a novel privacy model called AGeoI, based on the notion of GeoI by using a discrete road network graph. Our key contributions in this paper are outlined as follows.

- We present the notion of *approximate* geo-indistinguishability (AGeoI), a formal standard of location-privacy in a bounded co-domain, by generalising the classical paradigm of geo-indistinguishability We illustrate its applicability by proving the compositionality

theorem. Moreover, we show that the truncated Laplace mechanism canonically guarantees AGeoI.
- We propose a two-fold privacy-preserving navigation method for EVs dynamically querying for CS on road networks – the method protects against threats to individual locations of queries with formal AGeoI guarantees and against adversaries tracing the trajectories of the EVs in an online setting.
- Using real vehicular data and real locations of CSs from San Francisco, we experimentally show that our method ensures a very high fraction of EVs to have "privacy for free" and that the utility-loss for the EVs is very low compared to the gain in privacy.
- Our method not only ensures location-privacy guarantees but also enables EVs to estimate the real-time occupancy of CSs using sanitised queries to help users to plan their journeys efficiently.

The rest of this paper is organised as follows. Section II reviews some of the related work in this area. Section III introduces some fundamental notions on DP and GeoI. Section IV develops the mathematical theory of AGeoI. Section V elucidates the model of our proposed mechanism by formalizing the problem we are tackling, thoroughly discussing system architecture, and laying out the privacy-threat landscape we are addressing in this work. Section VI analyzes the cost of privacy on the EVs induced by our mechanism. Section VII presents the experimental results to illustrate the working of our mechanism, and Section VIII concludes the paper.

II. RELATED WORK

Both corporate and academic communities have recently piqued interest in advancing EVs and charging infrastructure to improve the transportation system's sustainability. Despite the advancements, the EV sector confronts challenges that delay the adoption process, such as range anxiety, an absence of convenient and available charging infrastructure and waiting time to charge [25], [26]. An offline static map of CS is insufficient to resolve these obstacles since EVs may need to reserve a charging station when a trip is planned or query the available stations based on their battery state, and CS must be reserved. Thus, live vehicular and charging station data is utilised in querying and reservation/scheduling mechanisms [27], [28], [29]. Encryption techniques can be used in such mechanisms to prevent external intrusions, but they cannot preserve users' privacy from malicious servers and third-parties.

Several data types are considered in these mechanisms, including real-time location, intended route, battery level, and station availability, to ensure the drivers are not detoured from their intended route [27], [30]. Although disclosing such information poses privacy concerns for the driver's location and vehicle tracking, the privacy requirements of such mechanisms are not sufficiently studied in the literature. Existing methods for planning charging points for EV journeys are considered mechanisms for confidentiality and integrity, but the drivers' location privacy is regarded as an issue of trust in the third-party service provider [31], [32].

This problem can be addressed by several approaches based on the threat model of the system. Location anonymity is achieved through cloaking an area [33], [34]. This approach can only be applied to the Edge of our system model to provide anonymity to a group of EVs, but we consider the Edge as an honest-but-curious threat actor and aim to preserve vehicles' privacy locally. Thus, such techniques are not trivially applicable to our considered threat model. Furthermore, anonymity techniques do not provide a formal privacy guarantee [35]. Similarly, mix-network approaches cannot be applied because there is no guarantee that multiple vehicles will be present in an Edge's coverage in any timestamp due to vehicles' movement [36].

An applicable approach to download the charging station's live map on EVs to search for the nearest or on-the-route available charging station has been considered and studied by the community [37]; however, the communication overhead of this technique is predicted to be much higher than the vehicles' location-based inquiry since it will require downloading a recent snapshot of the map for each query and, thus, has been criticised in the literature [38]. Moreover, due to the absence of data sharing, such methods hinder the statistical utility of the location data for the servers that may be useful for a variety of purposes (e.g. providing vital statistics to industries and institutions for optimally placing the CS on the map based on the query densities) and prevent the EVs from receiving any information about the traffic around and occupancy of certain CS restricting them to plan their journeys accordingly.

DP methods are gaining widespread usage in safeguarding location privacy across various domains, including automotive systems. The studies in [39], [40] proposed models by deploying a GeoI-based mechanism on the Edge for LBS. However, their approach did not consider preserving vehicles' locations against the Edge. An approach that complements the problem we aim to address in this paper was proposed by Qiu et al. in [41] where the authors proposed a technique to crowd-source a task in a vehicular network while preserving GeoI of the location of the vehicles offering Mobility as a Service in the spatial network to solve a task at a publicly known location in the map (e.g. taxi services). The problem formulation in this work is the inverse of what we aim to achieve in this paper. Hence, this work cannot be extended to address the privacy concerns induced by multiple dynamically generated queries throughout the journey.

In [42], Cunningham et al. studied the problem of trajectory sharing under DP and proposed a mechanism to address it. However, this work assumes the setting of an offline trajectory sharing which breaks down in the practical environment where the trajectories are being shared online as there is no prior information or limitation on the number of queries made by an EV during a journey and their respective locations. Therefore, the method proposed by the authors in [42] cannot be directly adapted to our dynamic environment closely simulating the real-world scenario for such a use case.

Of late, a major direction of research is along the lines of studying the statistical utility of differentially private data. A standard notion of statistical utility, which is extended to a variety of contexts, is the precision of the estimation of the distribution of the original data from that of the noisy data. Iterative Bayesian update (IBU) [43], [44], an instance of the famous *expectation maximization* method from statistics, provides one of the most flexible and powerful estimation techniques and has recently become in the focus of the community [45], [46]. In this work, we use IBU to approximate the distribution of the true locations of the queries made to the server and based on that, the users of the EVs can predict the availability of the CS around them in real-time and plan their route accordingly.

III. PRELIMINARIES

The most successful approach to formally address the privacy risks is DP, mathematically guaranteeing that the query output does not change significantly regardless of whether a specific personal record is in a dataset or not. Most research performed in this area probes two main directions. One is the classical central framework [13], [14], in which a trusted third-party (the curator) collects the users' personal data and obfuscates them with a differentially private mechanism.

Definition 3.1 (Differential privacy [13], [14]): For a certain query, a randomizing mechanism $\mathcal{R}$ provides ϵ-*DP* if, for all neighbouring[1] datasets, D and D', and all $S \subseteq \text{Range}(\mathcal{R})$, we have $\mathbb{P}[\mathcal{R}(D) \in S] \leq e^{\epsilon}\, \mathbb{P}[\mathcal{R}(D') \in S]$.

A major drawback of the central model is that it is vulnerable to security breaches because the entire original data is stored in a central server. Moreover, there is the risk that the curator may be corrupted. Therefore, a local variant of the central model has been widely popularized of late [15], where the users apply a randomizing mechanism locally on their data and send the perturbed data to the collector such that a particular value of a user's data does not have a major probabilistic impact on the outcome of the query.

Definition 3.2 (Local differential privacy [15]): Let $\mathcal{X}$ and $\mathcal{Y}$ denote the spaces of original and noisy data, respectively. A randomizing mechanism $\mathcal{R}$ provides ϵ-*LDP* if, for all $x,\ x' \in \mathcal{X}$, and all $y \in \mathcal{Y}$, we have $\mathbb{P}[\mathcal{R}(x) = y] \leq e^{\epsilon}\, \mathbb{P}[\mathcal{R}(x') = y]$.

Recently, GeoI [16], a variant of the local DP capturing the essence of the distance between locations [17], [18] has been in focus as a standard for privacy protection for location-based services, being motivated by the idea of preserving the best possible quality of service despite the local obfuscation operated on the data.

Definition 3.3 (Geo-indistinguishability [16]): Let $\mathcal{X}$ be a space of locations and let $d_{\text{E}}(x, x')$ denote the *Euclidean distance* between $x \in \mathcal{X}$ and $x' \in \mathcal{X}$. A randomizing mechanism $\mathcal{R}$ is ϵ-*geo-indistinguishable* if for all $x_1,\ x_2 \in \mathcal{X}$, and every $y \in \mathcal{X}$, we have $\mathbb{P}[\mathcal{R}(x) = y] \leq e^{\epsilon d_{\text{E}}(x_1, x_2)}\, \mathbb{P}[\mathcal{R}(x') = y]$.

Definition 3.4 (Iterative Bayesian update [43], [44]): Let $\mathcal{C}$ be a privacy mechanism that locally obfuscates points from the discrete space $\mathcal{X}$ to $\mathcal{Y}$ such that $\mathcal{C}_{xy} = \mathbb{P}(y|x)$ for all $x, y \in \mathcal{X}, \mathcal{Y}$. Let $X_1, \ldots, X_n$ be i.i.d. random variables on $\mathcal{X}$ following some distribution $\pi_{\mathcal{X}}$. Let Y_i denote the random variable of the output when X_i is obfuscated with $\mathcal{C}$.

Let $\mathrm{y} \in \mathcal{Y}^n$ be a realisation of $\{Y_1, \ldots, Y_n\}$ and q be the empirical distribution obtained by counting the frequencies of each element of $\mathcal{Y}$ as observed in y. The *iterative Bayesian update (IBU)* is a cutting-edge and strong technique for expectation maximization in statistics that can be used to estimate $\pi_{\mathcal{X}}$ by converging to the maximum likelihood estimate of $\pi_{\mathcal{X}}$ with the knowledge of y and $\mathcal{C}$. IBU works as follows:

1) Start with any full-support PMF θ_0 on $\mathcal{X}$.
2) Iterate $\theta_{t+1}(x) = \sum_{y \in \mathcal{Y}} q(y) \frac{\theta_t(x)\mathcal{C}_{xy}}{\sum_{z \in \mathcal{X}} \theta_t(z)\mathcal{C}_{zy}}$ for all $x \in \mathcal{X}$.

TABLE 1. List of Key Notations

Notation	Description
$\mathcal{X}$	Domain of original locations
$d_{\mathcal{X}}$	Distance on $\mathcal{X}$
$\mathcal{Y}$	Domain of obfuscated locations
$d_{\mathcal{Y}}$	Distance on $\mathcal{Y}$
$\mathbb{P}_{\mathcal{K}}[y\|x]$	Prob. that mechanism $\mathcal{K}$, applied to value x, reports y
I	Fixed Edge in the network
$R(I)$	Area of coverage by I
m	Number of locations reported by each EV
l_u	Vector of locations reported by EV u
$\mathcal{L}(t)$	Set of location vectors received by I at time t
$\mathcal{L}'(t)$	Shuffled set of all individual locations queried at time t
$\mathcal{R}(t)$	Set of nearest CS for $\mathcal{L}'(t)$
G	Road network graph
d_G	Travelling distance in graph G

IV. APPROXIMATE GEO-INDISTINGUISHABILITY (AGEOI)

In the classical framework of GeoI [16], the space of the noisy data is, in theory, unbounded under the planar Laplace mechanism. Under a certain level of GeoI that is achieved, the planar Laplace mechanism ensures a non-zero probability of obfuscating an original location to a privatised one which may be quite far, thus inducing a possibility of a substantial deterioration in the QoS of the users. This loss of QoS can be more sensitive in the context of the navigation of EVs, where it is extremely important to prioritize a bounded domain where a user is willing to drive – this may be a result of time constraints, the rising cost of fuel, geographical boundaries (e.g. international borders), etc. – giving rise to an idea of *area of interest* for each EV. This motivated us to extend the classical GeoI to a more generalized, approximate paradigm, inspired by the approach of the development of approximate DP from its pure counterpart.

Let $\mathcal{X}$ and $\mathcal{Y}$ be the spaces of the real and noisy locations equipped with distance metrics $d_{\mathcal{X}}$ and $d_{\mathcal{Y}}$, respectively. In general, $(\mathcal{X}, d_{\mathcal{X}})$ and $(\mathcal{Y}, d_{\mathcal{Y}})$ may be different and unrelated. However, for simplicity, here we assume $\mathcal{X} \subseteq \mathcal{Y}$ and, therefore, $d_{\mathcal{X}} = d_{\mathcal{Y}} = d$, and we proceed to define the notion of *approximate geo-indistinguishability*. It is worth noting here that, to an extent, we abuse the formal notion of "metric"

[1] Differing in exactly one place.

as d is not required to be symmetric, i.e., there may exist $x_1,\ x_2 \in \mathcal{Y}$ such that $d(x_1, x_2) \neq d(x_2, x_1)$.

Definition 4.1 (Approximate geo-indistinguishability): A mechanism $\mathcal{K}$ is *approximately geo-indistinguishable (AGeoI)* or (ϵ, δ)-*geo-indistinguishable* if for every measurable $S \subseteq \mathcal{Y}$, any pair of secrets $x, x' \in \mathcal{X}$, and for $\epsilon, \delta \in \mathbb{R}_{\geq 0}$ satisfying $\delta e^{\epsilon d(x,x')} \in [0, 1]$:

$$\mathbb{P}_{\mathcal{K}}\left[y \in S|x\right] \leq e^{\epsilon\, d(x,x')}\mathbb{P}_{\mathcal{K}}\left[y \in S|x'\right] + \delta\, e^{\epsilon d(x,x')} \tag{1}$$

One of the biggest advantages of DP and all of its variants that are accepted by the community is the property of compositionality, where the level of privacy can be formally derived with a repeated number of queries. Thus, we now enable ourselves to investigate the working of the compositionality theorem with the AGeoI which we defined, to stay consistent with the literature [35].

Theorem 4.1 [Compositionality Theorem for AGeoI]: Let mechanisms $\mathcal{K}_1$ and $\mathcal{K}_2$ be (ϵ_1, δ_1) and (ϵ_2, δ_2) geo-indistinguishable, respectively. Then their composition is $(\epsilon_1 + \epsilon_2,\ \delta_1 + \delta_2)$-geo-indistinguishable. In other words, for every $S_1,\ S_2 \subseteq \mathcal{Y}$ and all $x_1,\ x_1',\ x_2,\ x_2' \in \mathcal{X}$, we have:

$$\begin{aligned}&\mathbb{P}_{\mathcal{K}_1,\mathcal{K}_2}\left[(y_1, y_2) \in S_1 \times S_2|(x_1, x_2)\right]\\ &\leq e^{\epsilon_1\, d(x_1,x_1')+\epsilon_2\, d(x_2,x_2')}\\ &\mathbb{P}_{\mathcal{K}_1,\mathcal{K}_2}\left[(y_1, y_2) \in S_1 \times S_2|(x_1', x_2')\right]\\ &+ (\delta_1 + \delta_2)\, e^{d(x_1,x_1')+d(x_2,x_2')}\end{aligned} \tag{2}$$

Proof: Let us simplify the notation and denote:
$P_i = \mathbb{P}_{\mathcal{K}_i}[y_i \in S_i|x_i],\ P_i' = \mathbb{P}_{\mathcal{K}_i}[y_i \in S_i|x_i']$,
$\tilde{\delta}_i = \delta_i\, e^{d(x_i,x_i')}$ for $i \in \{1, 2\}$. As mechanisms $\mathcal{K}_1$ and $\mathcal{K}_2$ are applied independently, we have:

$$\mathbb{P}_{\mathcal{K}_1,\mathcal{K}_2}\left[(y_1, y_2) \in S_1 \times S_2|(x_1, x_2)\right] = P_1P_2 \tag{3}$$

$$\mathbb{P}_{\mathcal{K}_1,\mathcal{K}_2}\left[(y_1, y_2) \in S_1 \times S_2|(x_1', x_2')\right] = P_1'P_2' \tag{4}$$

Therefore, $\mathbb{P}_{\mathcal{K}_1,\mathcal{K}_2}[(y_1, y_2) \in S_1 \times S_2|(x_1, x_2)] = P_1P_2$

$$\begin{aligned}&\leq \left(\min\left(1 - \tilde{\delta}_1,\, e^{\epsilon_1\, d(x_1,x_1')}P_1'\right) + \tilde{\delta}_1\right)\\ &\quad\times \left(\min\left(1 - \tilde{\delta}_2,\, e^{\epsilon_2\, d(x_2,x_2')}P_2'\right) + \tilde{\delta}_2\right)\\ &\leq m_1m_2 + \tilde{\delta}_1m_2 + m_1\tilde{\delta}_2 + \tilde{\delta}_1\tilde{\delta}_2\\ &\quad\left[\text{where } m_i = \min\left(1 - \tilde{\delta}_i,\, e^{\epsilon_i\, d(x_i,x_i')}P_i'\right)\right]\\ &\leq e^{\epsilon_1\, d(x_1,x_1')+\epsilon_2\, d(x_2,x_2')}P_1'P_2'\\ &\quad+ \tilde{\delta}_1 - \tilde{\delta}_1\tilde{\delta}_2 + \tilde{\delta}_2 - \tilde{\delta}_1\tilde{\delta}_2 + \tilde{\delta}_1\tilde{\delta}_2\\ &\leq e^{\epsilon_1\, d(x_1,x_1')+\epsilon_2\, d(x_2,x_2')}\\ &\quad\times \mathbb{P}_{\mathcal{K}_1,\mathcal{K}_2}\left[(y_1, y_2) \in S_1 \times S_2|(x_1', x_2')\right]\\ &\quad+ (\delta_1 + \delta_2)\, e^{d(x_1,x_1')+d(x_2,x_2')}\end{aligned}$$

■

We now proceed to generalize the conventional planar Laplace mechanism [47] to define the *truncated Laplace mechanism* extended to a generic metric space.

Definition 4.2 (Truncated Laplace mechanism): The *truncated Laplace mechanism* $\mathcal{L}$ on a space $\mathcal{X}$ equipped with, not necessarily symmetric, distance metric d truncated to a radius r, is defined as:

$$\mathbb{P}_{\mathcal{L}}[y|x] = \begin{cases} c\, e^{-\epsilon\, d(y,x)} & \text{if } d(x, y) \leq r\\ 0 & \text{otherwise}\end{cases} \tag{5}$$

where c is the truncated normalization constant defined such that $\int_{y\in\mathcal{Y}} \mathbb{P}_{\mathcal{L}}[y|x]dy = 1$, and ϵ is the desired privacy parameter. Let us call r to be the *radius of truncation* for $\mathcal{L}$.

Note that for a discrete domain $\mathcal{Y}$, c is defined by normalizing $\sum_{y\in\mathcal{Y}} \mathbb{P}_{\mathcal{L}}[y|x] = 1$, and, in this case, $\mathcal{L}$ is a truncated geometric mechanism [48] extended to a generic metric space.

Lemma 4.2: For every $x_1, x_2 \in \mathcal{X}$ and $y \in \mathcal{Y}$, we have $e^{-\epsilon\, d(x_1,x_2)}\mathbb{P}_{\mathcal{L}}[y|x_1] - \mathbb{P}_{\mathcal{L}}[y|x_2] \leq 1$.

Proof:

$$\begin{aligned}&e^{-\epsilon\, d(x_1,x_2)}\mathbb{P}\left[y|x_1\right] - \mathbb{P}\left[y|x_2\right] \leq 1\\ &\iff c\left(e^{-\epsilon(d(x_1,x_2)+d(x_1,y))} - e^{-\epsilon d(x_2,y)}\right) \leq 1\end{aligned} \tag{6}$$

Now we observe that $d(x_1, x_2) + d(x_1, y) \geq d(x_2, y)$ due to the fact that d is a metric and it satisfies the triangle inequality. Immediately, we have $e^{-\epsilon d(x_1,x_2)+d(x_1,y)} - e^{-\epsilon d(x_2,y)} \leq 0$ for any $\epsilon \in \mathbb{R}_{\geq 0}$. Therefore, as $c \geq 0$, (6) is trivially satisfied. ■

Theorem 4.3: $\mathcal{L}$ satisfies (ϵ, δ)-geo-indistinguishability where

$$\delta = \max\left\{\max_{\substack{y\in\mathcal{Y}\\ x_1,x_2\in\mathcal{X}}} e^{-\epsilon\, d(x_1,x_2)}\mathbb{P}_{\mathcal{L}}\left[y|x_1\right] - \mathbb{P}_{\mathcal{L}}\left[y|x_2\right], 0\right\}.$$

Proof: Trivially $\delta e^{d(x_1,x_2)} > 0$ for any $x_1, x_2 \in \mathcal{X}$ as $\delta > 0$. Moreover, Lemma 4.2 ensures that $\delta e^{d(x_1,x_2)} < 1$. Now observe that for every $S \subseteq \mathcal{Y}$ and for all $x_1, x_2 \in \mathcal{X}$, we have:

$$\begin{aligned}&e^{-\epsilon\, d(x_1,x_2)}\mathbb{P}_{\mathcal{L}}\left[y|x_1\right] - \mathbb{P}_{\mathcal{L}}\left[y|x_2\right] \leq \delta\, e^{(1-\epsilon)\, d(x_1,x_2)}\\ &\Rightarrow \mathbb{P}_{\mathcal{L}}\left[y|x_1\right] - e^{\epsilon\, d(x_1,x_2)}\mathbb{P}_{\mathcal{L}}\left[y|x_2\right] \leq \delta\, e^{d(x_1,x_2)}\end{aligned}$$

■

The explicit process of sampling private locations satisfying AGeoI from a given set of original locations through a truncated Laplace mechanism on a discrete location space has been described in Algorithms 1 and 2.

V. SYSTEM MODEL

This section details our privacy-preserving model for finding an optimal CS in the Internet of Vehicles (IoV), as a use case of the proposed AGeoI technique. We begin with a discussion of the location privacy problems inherent in finding optimal CS in the IoV. This is followed by road networking modelling, a description of the system architecture for differentially private location sharing, the trust relationship between system tiers, and the privacy threat model.

Algorithm 1: Discrete and Truncated Laplace Mechanism (DTLap).

Input: Discrete domain of original locations: $\mathcal{X}$, Discrete domain of private locations: $\mathcal{Y}$, Desired privacy parameter: ϵ, Desired truncation radius: r;
Output: Channel C satisfying (5);
Function $\texttt{DTLap}(\mathcal{X}, \mathcal{Y}, \epsilon, r)$**:**
  Set $C \leftarrow$ empty channel;
  Set $Y \leftarrow$ empty list;
  for $x \in \mathcal{X}$ **do**
    $c_x = \frac{1}{\sum_{\substack{y \in \mathcal{Y} \\ d(x,y) \le r}} e^{-\epsilon\, d(x,y)}}$;
    for $y \in \mathcal{Y}$ **do**
      if $d(x, y) \le r$ **then**
        $C[x, y] = 0$
      else
        $C[x, y] = c_x\, e^{-\epsilon\, d(x,y)}$
  Return: C;

Algorithm 2: Sampling Private Locations With DTLap (DTLapSamp).

Input: Discrete domain of original locations: $\mathcal{X}$, Discrete domain of private locations: $\mathcal{Y}$, Desired privacy parameter: ϵ, Desired truncation radius: r; Vector of original locations: X;
Output: Corresponding vector of private locations: Y;
Function $\texttt{DTLapSamp}(\mathcal{X}, \mathcal{Y}, \epsilon, r, X)$**:**
  $C = \text{DTLAP}(\mathcal{X}, \mathcal{Y}, \epsilon, r)$;
  Set $Y \leftarrow$ empty list;
  for $x \in X$ **do**
    Randomly sample $y \in \mathcal{Y} \sim C[x, :]$;
    Append y to Y
  Return: Y;

A. PROBLEM STATEMENT

EVs have emerged as crucial components of future sustainable transportation systems, aimed at reducing CO2 emissions. Consequently, they have received considerable attention from both academia and industry [26]. However, due to their limited battery capacity, EVs often need to visit CS during journeys. This requirement leads to range anxiety among some drivers, where they fear that their vehicles lack sufficient battery power to reach their intended destinations. Range anxiety is recognized as a major obstacle to the broad acceptance of EVs [49]. While CS are not always readily available, as it takes time to sufficiently charge EVs, the implementation of a CS booking service can help alleviate range anxiety.

To minimize charging wait times, EVs can access CS booking services through third-party providers, enabling them to discover the nearest and readily available CS. This can be achieved through static or live location queries. However, location sharing raises privacy challenges, such as vehicle tracking. GeoI technique provides a formal privacy guarantee for location queries. However, it is not highly applicable to this use case for two reasons. It does not consider the feasible locations where a vehicle can be present, and it does not stop vehicle tracking in the case of linked queries during the vehicle trajectory. Thus, a tailored privacy-preserving mechanism is facilitated by combining the proposed AGeoI technique with dummy location generation.

B. ROAD NETWORK MODEL

Similar to [41], the road network G is represented as a weighted directed graph $G = (N, E, W)$, where N is the set of nodes, $E \subseteq N^2$ is the set of edges, and $W : N^2 \to \mathbb{R}^+$ is the set of weights representing the minimum travelling distance between any two nodes. The nodes and edges correspond to junctions and road segments of the network, respectively. Each edge $e \in E$ is addressed by the pair of respective starting node, ending node, and a weight representing the travelling distance through that edge, i.e., $e = (N_e^s, N_e^e, w_e) \in N$, where the direction of the traffic is from N_e^s to N_e^e on e. For any $i \in N$ and $j \in N$, let the sequence of edges $(e_1, \ldots, e_r)$ denote a *path* from node i to node j if $N_{e_1}^s = i$ and $N_{e_r}^e = j$. Hence, let $C(i, j)$ represent the set of paths that connect node i to node j. Then W is a $N \times N$ matrix, where

$$W_{ij} = \begin{cases} \min\limits_{p \in C(i,j)} \sum\limits_{e \in p} w_e & \text{if } C(i, j) \neq \phi \\ \infty & \text{otherwise} \end{cases}$$

Essentially W_{ij} is the shortest travelling distance from node i to node j in the network. We shall address the quantity W_{ij} as the *traversal distance* between nodes i and j in the graph G and denote it as $d_G(i, j)$ for every $(i, j) \in N^2$. Note that, as G is a directed graph, d_G may not be symmetric.

C. SYSTEM ARCHITECTURE

IoV applications are revolutionising transportation systems by mitigating human errors, enhancing travel convenience, and reducing energy, operational, and environmental costs [50], [51]. EVs have emerged as a viable technology for lowering carbon emissions and travel costs [52]. However, range anxiety is one of the major challenges of their wide adoption. Vehicular location data can be utilised to optimise the vehicle charging plan and mitigate range anxiety. Third-party services can assist users by recommending available CS in close proximity. However, depending on these third-party providers gives rise to notable privacy concerns within the threat model of honest-but-curious service providers, which in turn requires users to place their trust in them.

The system architecture, illustrated in Fig. 1, incorporates vehicles within an ITS that operates on a three-tier architecture. This architecture comprises Roadside Units (RSUs) connected to a Mobile Edge Computing (MEC) Server, which is connected to the Core Cloud through a secure communication channel. The Core Cloud facilitates the connection between

vehicles and third-party services, including the charging station recommender system, which is the main focus of this paper. However, guaranteeing the complete trustworthiness of the cloud architecture and third-party service providers in handling vehicular location data is not feasible, aligning with the honest-but-curious threat model. Consequently, our proposed architecture only shares privatised vehicular location data. The subsequent sections delve into a comprehensive description of the roles and functions of each system component.

1) VEHICLE TIER

We fix a road network G with nodes $G(N)$ and edges $G(E)$. We choose an arbitrary edge $I \in G$, and focus on the queries made by the EVs in I's range of coverage, $R(I)$, provided by its RSU. An EV u employs a local obfuscation technique to protect its true location $x^u \in R(I)$ to $x_1^u \in R(I)$ within the coverage area $R(I)$ of a specific edge. When an EV moves from the area of coverage of one Edge cloud to another, we can assume the queries and the privacy threats against the Edge to reset as each Edge communicates with the Cloud-based services and the third-party service providers.

The vehicle u utilises DTLap and DTLapSamp algorithms to apply the truncated Laplace mechanism, guaranteeing AGeoI. DTLap creates a probabilistic mapping from each original location to a set of private locations, ensuring that each mapping adheres to the differential privacy constraints specified by ϵ. The truncation radius r limits how far a private location can be from the original location, enhancing practical utility. Using the channel created by DTLap, DTLapSamp generates a vector of private locations that correspond to a given vector of original locations. The result is a vector of locations that preserves privacy while reflecting the distribution of the original locations. Then the vehicle generates $m-1$ plausible dummy locations $\{x_2^u, \ldots, x_m^u\} \in R(I)^{m-1}$ in the coverage area of the respective edge with the privatised location and reports the vector of m locations, $l_u = (x_1^u, \ldots, x_m^u)$, to I. At any given time, u locally obfuscates its true location $x^u \in R(I)$ to $x_1^u \in R(I)$ using a truncated Laplace mechanism guaranteeing AGeoI and generates $m-1$ feasible dummy locations $\{x_2^u, \ldots, x_m^u\} \in R(I)^{m-1}$ in the coverage area of the respective edge. Then u reports the vector of m locations, $l_u = (x_1^u, \ldots, x_m^u)$, to I for the Edge to process and communicate the query to the cloud services and the third parties to find the nearest available CS in $R(I)$.

2) EDGE TIER

Given the substantial volume of data generated and exchanged between vehicles and infrastructure, the installation of edge clouds in close proximity to vehicles becomes essential to host off-board vehicular services, which require low access latency from onboard vehicular services [53]. In addition to performing essential data processing and forwarding functions, the Edge also serves as a layer for data aggregation. Moreover, it enables the deployment of supplementary privacy-preserving measures before sharing the data with third-party entities.

3) CLOUD TIER

It is expected to provide computation and storage capabilities for top-level processes, including data-sharing interfaces for third-party services.

4) THIRD-PARTY SERVICE PROVIDER

It is the external party to ITS and is expected to enhance the quality of the function for finding the available CS for the vehicles by receiving search queries compromised of privatised and dummy location vectors for the respective vehicles.

5) COMMUNICATION CHANNEL

ITS comprises a network of RSU, vehicle on-board electronic control units (ECU), and distributed cloud computing and storage services. Wireless communications are enabled for V2V (Vehicle to Vehicle), V2I (Vehicle to Infrastructure) and V2X (Vehicle to Everything), facilitated by the technologies such as IEEE 802.11p DSRC/WAVE (Dedicated Short Range Communication/Wireless Access in Vehicular Environments), cellular advances such as C-V2X, and the long-term evolution for vehicles (LTE-V) [54]. Confidentiality of the wireless communication channel is secured by public key infrastructure (PKI) encryption methods which are beyond the scope of this work.

D. PRIVACY THREAT LANDSCAPE

In real-time IoV location-based applications, it is often necessary for users to share their location information with the service provider in order to access location-specific services. However, this raises privacy concerns as it can potentially expose sensitive information about individuals' movements and activities. To address these concerns, data perturbation techniques can be employed to protect the privacy of users while still allowing them to access the services they need. These techniques introduce uncertainty or noise into the data, preventing an attacker from identifying the precise location of an individual. However, real-world solutions often rely on user consent, access control, and non-disclosure agreement-based mechanisms instead of providing formal privacy guarantees. Thus, there are existing privacy challenges related to shared location data, including journey tracing and location identification.

Furthermore, apart from these major privacy challenges, vehicular location data may also be susceptible to unauthorised use, data inference, retention, or insider privacy breaches within the service provider when formal privacy guarantees are lacking. The third-party provider is typically considered an honest-but-curious adversary model, assuming it is honest in accurately executing the protocol required to provide location data. However, there is a possibility that the provider may be curious about inferring users' private information based on the acquired location data [55].

This study aims to offer a formal privacy guarantee for location-based querying that can be utilised by vehicles throughout their trajectories to effectively address the associated privacy challenges with this process. To achieve this, the system is considered in three categories: (i) the vehicle users (data subject), (ii) ITS encompassing the Edge and Core Cloud Tiers (data controller and data processor), and (iii) the third-party that receives the privatised data through the deployed privacy-preserving mechanisms. The third-party is assumed to be an EV charging management system, which may operate under a registration-based approach for a specific area. Our focus is on mitigating the following two major sources of threats that have the potential to compromise the privacy of EVs.

1) LOCATION IDENTIFICATION

It is vital to ensure that the privatized version of the true location of the EV is within a certain radius of interest w.p. 1, making sure that the reported location is within a feasible and drivable distance away, and most importantly, within the area of coverage of the Edge where its true location lies. Therefore, we defined AGeoI as an extension of GeoI. Thus, to ensure the privacy of any given query in the road network, the EVs locally obfuscate their true locations using the truncated Laplace mechanism with their desired parameter ϵ and the radius of truncation r, which, in turn, decide the value of δ.

2) JOURNEY TRACING

EVs may inquire about the nearest available charging station, without proceeding with the query, and raise further queries along the journey. Subsequently, additional queries may be raised at different points during the journey. In our model, we aim to capture this realistic setting by allowing multiple queries to be made by the EV within a single journey. However, this introduces a potential threat of approximately tracing the trajectory of the EV's journey by interpolating the locations of the queries, despite each individual location being AGeoI-protected. This is due to the fact that the obfuscated location of each query is not distinguishable from the real location, but they are not too far off from each other with a very high probability. Consequently, if a large number of queries are made within a single journey, it becomes relatively straightforward to approximate the trajectory of the EV's journey.

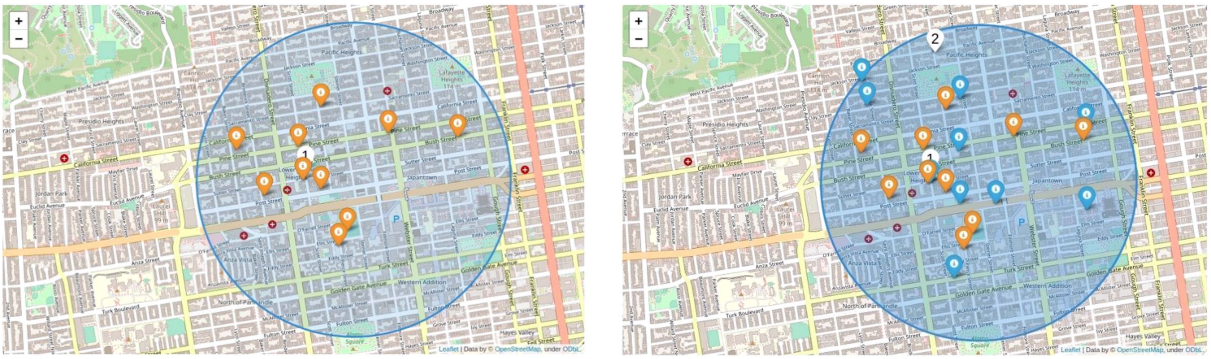
Cunningham et al. [42] proposed a mechanism to securely share trajectories under LDP. However, the authors in [42] assumed a model of offline sharing of the entire trajectory and, hence, sanitising it with the proposed mechanism to engender LDP guarantees. In our setting, this method cannot be directly implemented as we consider a dynamic environment where the queries made by the EVs are in real-time, with the server not having any prior knowledge of the number or the location of the queries made by a certain EV. Therefore, the mechanism of [42] cannot trivially be extended in the online location-sharing environment, and hence, the threat of adversaries able to reconstruct the journey of a particular EV with a high number of queries remains as a concern.

E. PROPOSED QUERY MODEL

During the journey, an EV u located within the coverage of an Edge I locally obfuscates its true location $x^u \in R(I)$ to $x_1^u \in R(I)$ using the truncated Laplace mechanism guaranteeing AGeoI, and generates $m-1$ feasible dummy locations $\{x_2^u, \ldots, x_m^u\} \in R(I)^{m-1}$, i.e., locations that cannot be trivially identified as being artificially generated given the query of the previous time stamp w.r.t. realistic speed limits, travelling conditions, etc. For the first query that u makes along its journey, it generates random $m-1$ dummy locations in $R(I)$. Thus, each query of u consists of reporting the vector of m locations, $l_u = (x_1^u, \ldots, x_m^u) \in R(I)^m$, to I for the Edge to process and communicate the query to the Cloud services and the third-parties to find the nearest available CS in $R(I)$. This approach ensures that the adversary will have at least m possible trajectories that the EV could have realistically followed at every time stamp, making it highly improbable for the Edge and the third-party to be able to conclude which of them was the actual journey as, after k queries made along a single journey, each interpolated trajectory will have a probability of at least $1/m^k$ of being the real one.

Fig. 2 illustrates 10 reported dummy locations along with the privatized location for two consecutive time windows. Notably, the dummy locations in the subsequent time window can be feasibly linked to at least one of the preceding dummy locations. At any given time, the Edge collects all the reported locations from the querying EVs, shuffles them by effacing the links between the location vectors and the corresponding EVs, and sends this jumbled collection of all the reported locations in the network to know their respective nearest available CS to the third-party service provider. After receiving the response, the Edge, which internally keeps the record of the IDs of the EVs against their queried locations, assigns the corresponding vector of locations of the nearest available CS to each EV and communicates them back to the respective vehicles.

In other words, at time t, if the Edge receives the location vectors from k_t querying EVs as $\mathcal{L}(t) = \{l_{u_1}, \ldots, l_{u_{k_t}}\}$, the Edge is responsible for shuffling all the individual locations in these reported vectors and forward the scrambled collection $\mathcal{L}'(t) = \{x_i^u : u \in \{u_1, \ldots, u_{k_t}, i \in [m]\}$ to the Cloud and the third-party, while internally keeping a track of the IDs of the EVs to reconnect the query-response back to the corresponding users. Setting $\hat{x}$ as the location of the nearest available charging station from location x in $R(I)$, the Edge receives $\mathcal{R}(t) = \{\hat{x}_i^u : u \in \{u_1, \ldots, u_{k_t}, i \in [m]\}$ as the response from the third-party service provider handling the CS data real-time. After this, matching the IDs of the EVs with the locations of the CS, the Edge communicates the response vector $\hat{l}_u = (\hat{x}_i^u : i \in [m])$ back to the corresponding EV u. Then the EV can choose to navigate to $\arg\min_{x \in \hat{l}_u}\{d_G(x, x_u)\}$, where x_u is the real location of u. The overview of this mechanism is given in Fig. 1.



(1) (2)

FIGURE 2. Reported dummy and privatised locations for two respective time windows (White Pins: Privatised locations, Orange Pins: Dummy locations in 1st Time window, Blue Pins: Dummy locations in 2nd Time window).

VI. COST OF PRIVACY ANALYSIS

Definition 6.1 (Cost of privacy): Suppose an EV u at location x^u chooses to locally obfuscate its real location of a query as x_1^u using the truncated Laplace mechanism $\mathcal{L}_{\epsilon,r}$ satisfying (ϵ, δ)-geo-indistinguishability with a corresponding radius of truncation r. Then we define the *cost of privacy (CoP)* of EV u as $\text{CoP}(u, \mathcal{L}_{\epsilon,r}) = \text{c}(x^u, \hat{x}_1^u) - \text{c}(x^u, \hat{x}^u)$, where $\hat{x}^u$ and $\hat{x}_1^u$ are the nearest available CS in the network to x^u and x_1^u, respectively, and $\text{c} : G(N)^2 \mapsto \mathbb{R}^+$ is any cost function that reflects the "cost" of the commute from locations x to y in the network.

In other words, CoP, as in Definition 6.1, essentially captures the *extra* cost that an EV needs to cover as a result of the privatized location it reports to the Edge satisfying AGeoI, as opposed to its true location. In this paper, for the purpose of simplicity of the analysis, we considered the cost function as the travelling distance in the network, i.e., $\text{c} = d_G$. However, in practice, any suitable cost function could be used (e.g. fuel efficiency, time, etc.) could be used as c, depending on the context and requirement of the architecture. To formally characterize and analyze the CoP of the EVs in the network, inspired from the classical version of *Voronoi decomposition*, we extend the concept in the setting of our road network in the network coverage for a fixed Edge w.r.t. graph-traversal distance, d_g.

Definition 6.2 (Voronoi decomposition): Let G be the graph representing the road network equipped with travelling distance d_G. Let the set of CS in G be $C_G = \{c_1, \ldots, c_{n_G}\}$. Then the *Voronoi decomposition* on G w.r.t. C_G is defined as $\text{V}_G = \{V_i : i \in [n_G]\}$ such that $V_i \cap V_j = \phi$ for any $i \neq j$ and $\bigcup_{i\in[n_G]} V_i = G$, where

$$V_i = \{x \in G : d_G(x, c_i) \leq d_G(x, c_j) \, \forall \, j \in [n_G], j \neq i\}$$

Definition 6.3 (Closed ball around a location): For any $x \in G$ and $r \in \mathbb{R}_{\geq 0}$, the *closed ball* of x of radius r is defined as $\beta_r(x) = \{y \in G : d_G(x, y) \leq r\}$

Definition 6.4 (Fenced Voronoi decomposition): For any $r \in \mathbb{R}_{\geq 0}$ and charging station i, let the *r-fenced Voronoi decomposition* on road network G be defined as $V_G^{-r} = \{V_i^r : i \in [n_G]\}$ such that $V_i^{-r} \cap V_j^{-r} = \phi$ for $i \neq j$ and $V_i^{-r} = \{x \in V_i : B_r(x) \subseteq V_i\}$. In other words, V_i^{-r} essentially constructs an area contained within V_i restricted by a *fence* at a distance r from the edge of V_i.

Theorem 6.1: Suppose an EV u positioned at x^u on G obfusucates its location using AGeoI with any radius of truncation $r \in \mathbb{R}_{\geq 0}$. Let $\hat{x}^u$ be the location of the nearest available charging station to the true location x^u. Then $\mathbb{P}[\text{CoP}(u, \mathcal{L}_{\epsilon,r}) = 0] = 1$ for every $x^u \in V_{\hat{x}^u}^{-r}$. In other words, if an EV lies in the r-fenced Voronoi decomposition for its nearest available CS, it has a *zero cost for privacy* w.p. 1.

Proof: Immediate from Definition 6.4. ■

Theorem 6.2: Suppose an EV u lies in $V_{\hat{x}^u} \setminus V_{\hat{x}^u}^{-r}$ and it uses AGeoI to obfuscate its true location x^u to x_1^u with a radius of truncation r and privacy parameter ϵ for making a private query to the Edge. Then $\mathbb{P}[\text{CoP}(u, \mathcal{L}_{\epsilon,r}) = 0] = 1 - \sum_{x_1^u \in V_{\hat{x}^u}^c} ce^{-\epsilon d_G(x^u, x_1^u)}$, where c is the normalizing constant of the truncated Laplace mechanism as in Definition 4.2.

Proof: To compute $\mathbb{P}[\text{CoP}(u, \mathcal{L}_{\epsilon,r}) = 0]$, we only need to exclude the possibilities where the reported location of the EV lies outside the Voronoi decomposition of the station $\hat{x}^u$, which, essentially, is $1 - \sum_{x_1^u \in V_{\hat{x}^u}^c} ce^{-\epsilon d_G(x^u, x_1^u)}$. ■

VII. EXPERIMENTAL STUDY

This section presents the experimental study with the objectives as follows: (i) to validate proposed theoretical claims and solutions empirically; (ii) to use the method to find the nearest available charging station for EVs as a case study; (iii) to investigate the cost of privacy in real-time settings; and (iv) to conduct a real-time CS occupancy prediction technique from the noisy vehicle distribution. Standard Python packages are used to run the experiments in an environment with an Intel core i7 processor, 16 GB of RAM, and an Ubuntu 20.04 operating system.

A. DATASET PREPARATION

The road network data extracted from OpenStreetMap [56]. The cost of privacy is calculated as the additional routing

distance caused by noise in vehicular locations during queries to identify the optimal charging station. The cost of privacy depends on the sparsity of CS. We prepared two datasets: one with 404 existing charging station locations in San Francisco obtained from the United States Department of Energy [57], and another by merging existing and planned charging station locations with on-street and off-street parking locations from DataSF [58], resulting in 716 independently distributed locations.

The EPFL mobility dataset includes GPS records of 536 taxi trajectories in San Francisco over four weeks [59]. The dataset provides information such as the taxi identifier, latitude, longitude, occupancy state (vacant or occupied), and a UNIX epoch timestamp. Leveraging the occupancy information, we were able to split the complete taxi trajectories into individual customer trajectories, resulting in over 450,000 exported trajectories. For our study, we randomly selected 536 trajectories from each taxi.

B. EXPERIMENTAL SETUP

A group of EVs sends out location queries to find the closest available CS during their journeys on the road network G. The edges of the road network G are truncated into discrete segments with an equal k travel distance, similar to the work in [41]. DTLap is utilised to generate the privacy channel by using the Laplace mechanism for the user's desired values for privacy budget ϵ and truncation radius r. Following this, DTLapSamp is used to generate privatised locations with respect to the users' real locations.

A location query contains a privatised location and $m-1$ dummy locations as a vector and is collected by the Edge for sending them to the third-party through the core cloud as a single vector of all locations. The third-party responds to the locations in the vector with the closest available CS for each, and the Edge sends vehicle location vectors to the related vehicles without being able to differentiate privatised and dummy locations.

For IBU to approximate the original distribution of the query locations of the EVs in the road network in order to predict the availability of the CS and, thus, assist the users in planning their journeys appropriately, we note that each original query location goes through two independent steps of sanitization: a) locally using the truncated Laplace mechanism to achieve AGeoI and b) generating the realistic dummy locations in the area of coverage of the Edge to ensure protection against attacks reconstructing their journeys. Setting the domain $\mathcal{X}$ as the area of coverage of the RSU of the fixed Edge that we focus on, while the former is a straightforward use of the channel $\mathcal{L}$, the latter can be thought of as $m-1$ independent applications of the uniform channel $\mathcal{U}$, where $U : \mathcal{X}^2 \mapsto \mathbb{R}$ with $\mathcal{U}_{x,y}$ denoting $\mathbb{P}_{\mathcal{U}}[y|x] = 1/|\mathcal{X}|$, by each EV. Therefore, after accounting for the normalization, the channel incorporating the local obfuscation and the generation of the dummy locations used by each EV reduces down to $\frac{1}{m}\mathcal{L} + \frac{m-1}{m}\mathcal{U}$ which we use as the privacy channel to implement IBU.

The first set of experiments examines the CoP for randomly selected 536 vehicle traces, where each trace contains a series of GPS coordinates and 3 randomly selected points along each for the real locations of the queries. The discrete road network is generated by setting the distance $k = 100$ meters. The parameters of ϵ and r are varied in the range of 0.2 to 2, and 1 to 20, respectively.

The privatised location, together with the dummy locations, is sent to the third-party for a query to prevent the third-party from tracking the vehicle. The area of Edge coverage, rather than the vehicle's area of interest, is considered for dummy location generation rather than the vehicle's area of interest, as the centre of mass may give away the true location. The second set of experiments examined the impact of dummy locations on the CoP.

The location queries could be used for real-time predictive analysis on the optimisation of the smart power grid, managing staff, and determining where new CS should be deployed. Thus, service providers can have the utility of the datasets (e.g., training ML models, etc.) with DP-based methods while the privacy of individuals is preserved. The third set of experiments utilises the IBU method to retrieve the true distribution of locations of the queries from the noisy distribution, which includes privatised and dummy locations.

C. RESULTS AND DISCUSSION

1) COST OF PRIVACY

DP approaches introduce a trade-off between privacy and data utility, with a higher level of privacy requiring a greater level of noise. The efficacy of the respective service may correspondingly decrease due to the fall in data utility, and this difference in the *quality of service* is referred to as the 'cost of privacy' (CoP) in this study. In particular, in the context of the use case considered in this paper, the CoP is formalised in Definition 6.1.

The following results are achieved by carrying out the experiments for 3 linked queries of 536 randomly selected vehicle trajectories for varying values of ϵ or r ranging from 0.2 to 2, and 1 to 20, respectively. Fig. 3 demonstrates the CoP in terms of the extra travelling distance due to the privacy-preserving mechanism, where a similar pattern is observed for both of the datasets. Another observation is that a high frequency of queries resulted in no cost for privacy preservation. Fig. 4 shows the fraction of the queries with "privacy for free" where both datasets followed similar patterns. Vehicle queries contain dummy locations and their privatised true locations. It is possible that the dummy locations can sometimes provide a better utility, but our experiments consider the utility of a privatised location as the worst-case for analysis.

Fig. 3 shows that our method provides a negligible cost of utility loss for the formal privacy gain enjoyed by the EVs. By increasing the truncation radius, an abrupt drop in the distance between the location of the nearest available charging station for the true location of the query and that of the privatised one

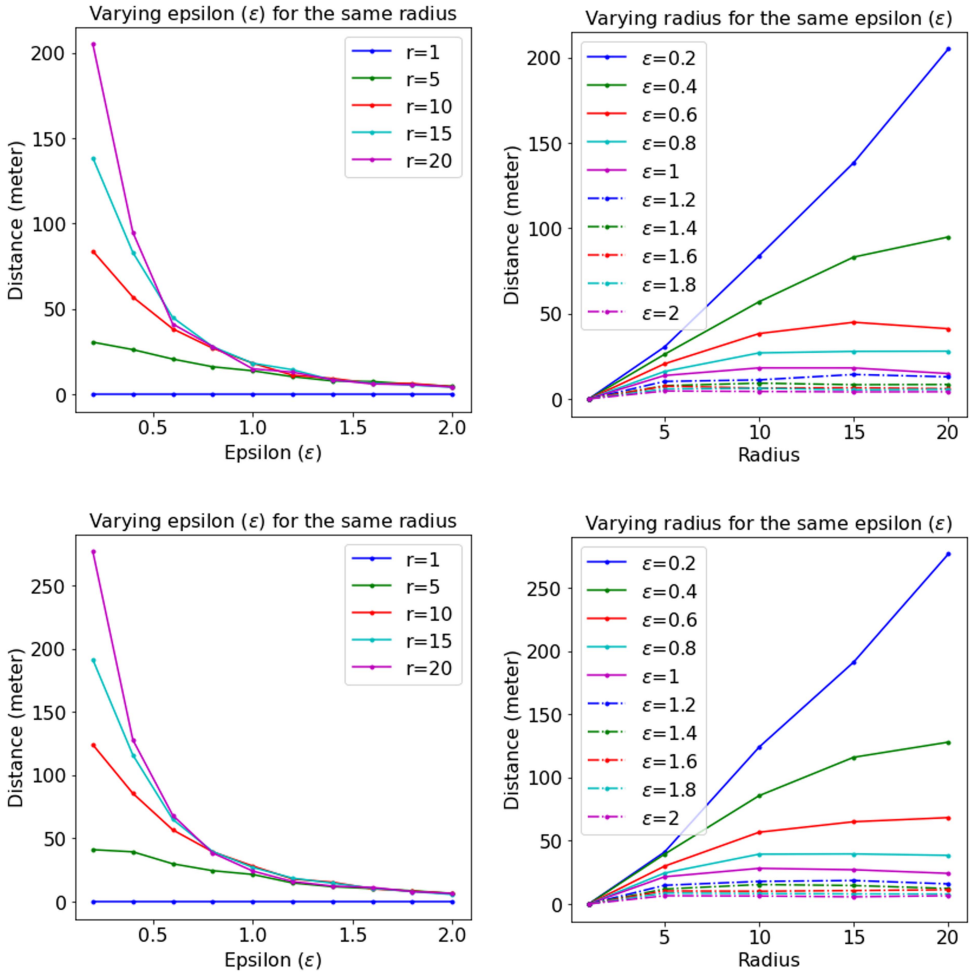


FIGURE 3. CoP (i.e., by Definition 6.1, the difference in the distance an EV needs to cover to reach the nearest CS with and without local obfuscation to achieve AGeoI) for varying ϵ or r of AGeoI (1st row is for sparse CS, 2nd row is for dense CS).

implies that the cost of the extra travel distance needed to be taken due to the AGeoI guarantee is almost negligible. A similar trend is seen for the varying ϵ with a fixed radius. As the level of privacy decreases, the fraction of EVs in the network enjoying *privacy for free* grows to be more than 60% for a radius of truncation of merely 10 road segments, where each segment is 100 meters long, for $\epsilon \geq 0.5$. However, more than 90% of the EVs achieve a zero cost of privacy for $\epsilon \geq 1.5$, irrespective of the truncation radius as illustrated in Fig. 4. Due to increasing perturbation for the disclosed locations, the width of the confidence interval for zero cost of privacy increases, as seen in Fig. 4. The likelihood of achieving zero cost of privacy fluctuates over a wider range and it does not monotonically decrease with the growing radius due to rising randomness.

2) IMPACT OF DUMMY DATA GENERATION

Considering an adversary interested in finding the true locations of the EVs, (α, β)-*identifiability* is defined for any location x as $\mathbb{P}[d(y, x) < \alpha) > \beta$, where y is any guessed location by the adversary. With the proposed method, with a sufficiently small radius of truncation to obfuscate the true location using the truncated Laplace and generating $m-1$ dummy locations in the area of coverage of the Edge, the probability of hitting the true x within an error of α is $\mathbb{P}[d(x, y) < \alpha] = m^{-1}ce^{-\epsilon\alpha} = \beta$, where c is the normalising constant.

There has been some work in this area from the perspective of just GeoI [39], [40], [41], [60], [61] or just from the standpoint of generating dummy locations exploiting anonymisation techniques [62], [63]. One of the first major concerns in using only GeoI is when we allow dynamic and multiple queries along the journey of the EVs, as individual locations, despite being privatised, can still be interpolated to approximate the entire trace. If only dummy locations are used, however, any estimated (or observed) y could be the real location w.p. $\frac{1}{m-1}$, as there is no formal privacy guaranteed, i.e., every location x has, is $(0, (m-1)^{-1})$-identifiable among $(m-1)$ dummy locations. With potential parallel processing, brute-force attacks are just one way that it has been shown

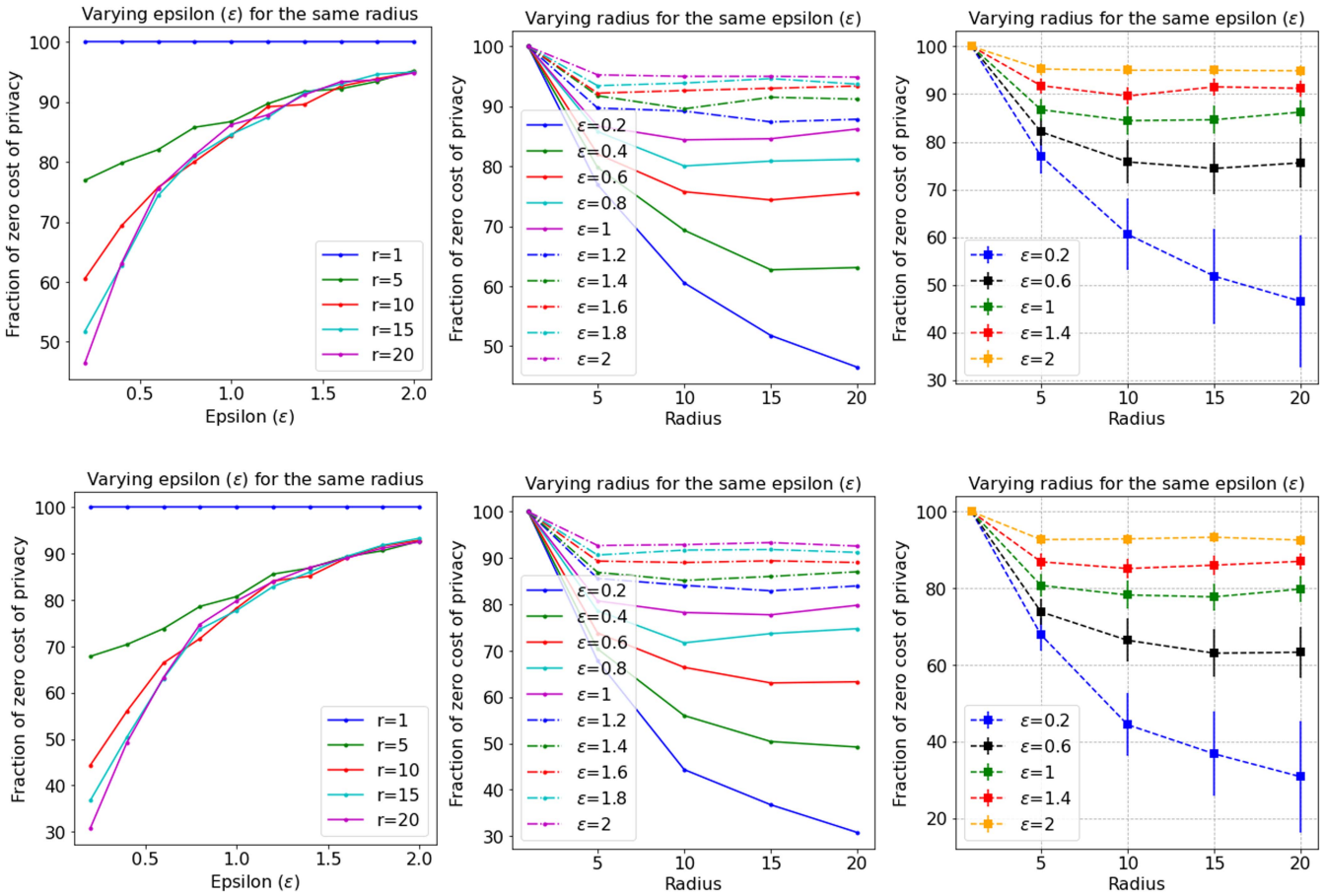


FIGURE 4. Fraction of EVs incurring no CoP for varying ϵ or r of AGeoI (1st row is for sparse CS, 2nd row is for dense CS).

that anonymisation techniques are not sufficient to protect privacy [64].

Fig. 5 illustrates how the CoP increases with an increase in the noise due to the lack of dummy locations under the same level of identifiability. To achieve the same (α, β)-identifiability with just AGeoI without dummy locations, the parameter ϵ needs to be scaled by $(\ln m)^{-1}$, i.e., more noise needs to be added, which results in having a worse trade-off between privacy and CoP for the same level of privacy.

3) REAL-TIME PREDICTIVE STUDY

Predicting the availability of CS is a vital aspect of EV trip planning, offering a solution to alleviate range anxiety. Existing methodologies predominantly leverage machine learning-based approaches for such predictions [65], [66], [67], [68], [69]. These models generally focus on pre-booking of CS timeslots, drawing upon historical data such as previous CS usage, traffic density, and external factors like weather conditions. However, the static nature of these predictions may not fully accommodate the dynamic and often unpredictable fluctuations in traffic patterns. The rigid scheduling could lead to scenarios where an EV misses its charging slot due to traffic delays or a CS remains unutilized despite being available. Hence, a real-time predictive analysis would be critical to determine the likelihood of a CS being available when an EV arrives.

To address this gap, our proposed method emphasises real-time predictive analysis, crucial for assessing CS availability upon an EV's arrival. By preserving privacy and capturing live traffic distribution data from querying vehicles, our approach utilises IBU to estimate current traffic conditions. The statistical distance between the estimated and the original distributions are shown in Fig. 6. We considered two different levels of AGeoI with $\epsilon = 0.6$ and $\epsilon = 2$ and IBU was run for 100 iterations. The results demonstrate that the distance between the original and the estimated distributions of the traffic is decreasing. The accuracy of the estimation of the original distribution from the noisy locations is illustrated by the heatmaps of Fig. 6 depicting the original, noisy, and estimated traffic distributions. This essentially highlights the high statistical utility of our proposed method and, specifically, helps in the prediction of how likely a CS will be available when the vehicle arrives and the traffic, in general.

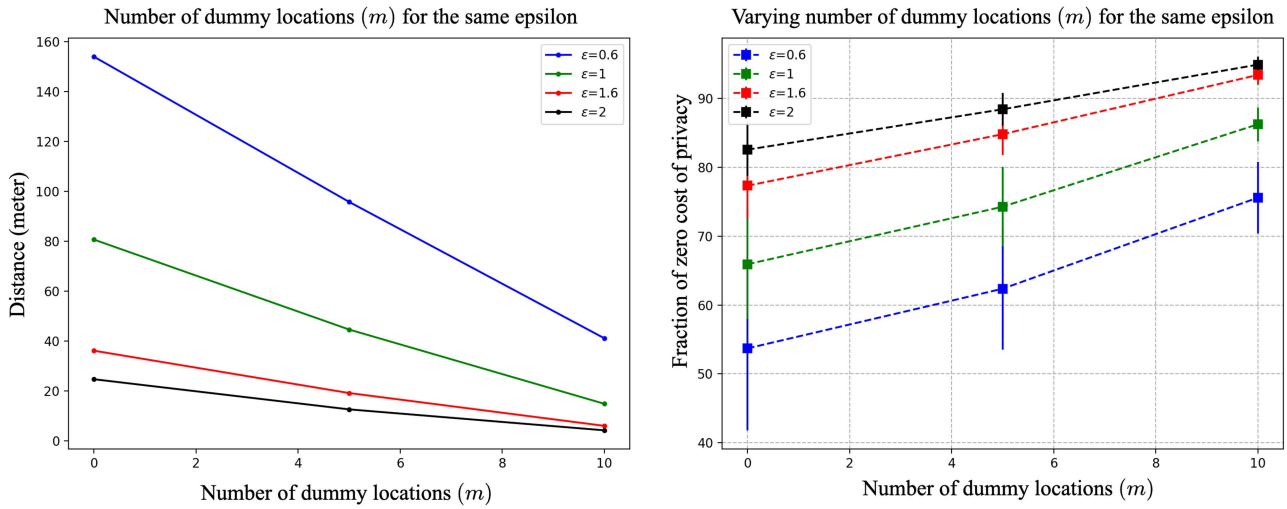


FIGURE 5. Impact of introducing dummy locations along with AGeoI on the CoP.

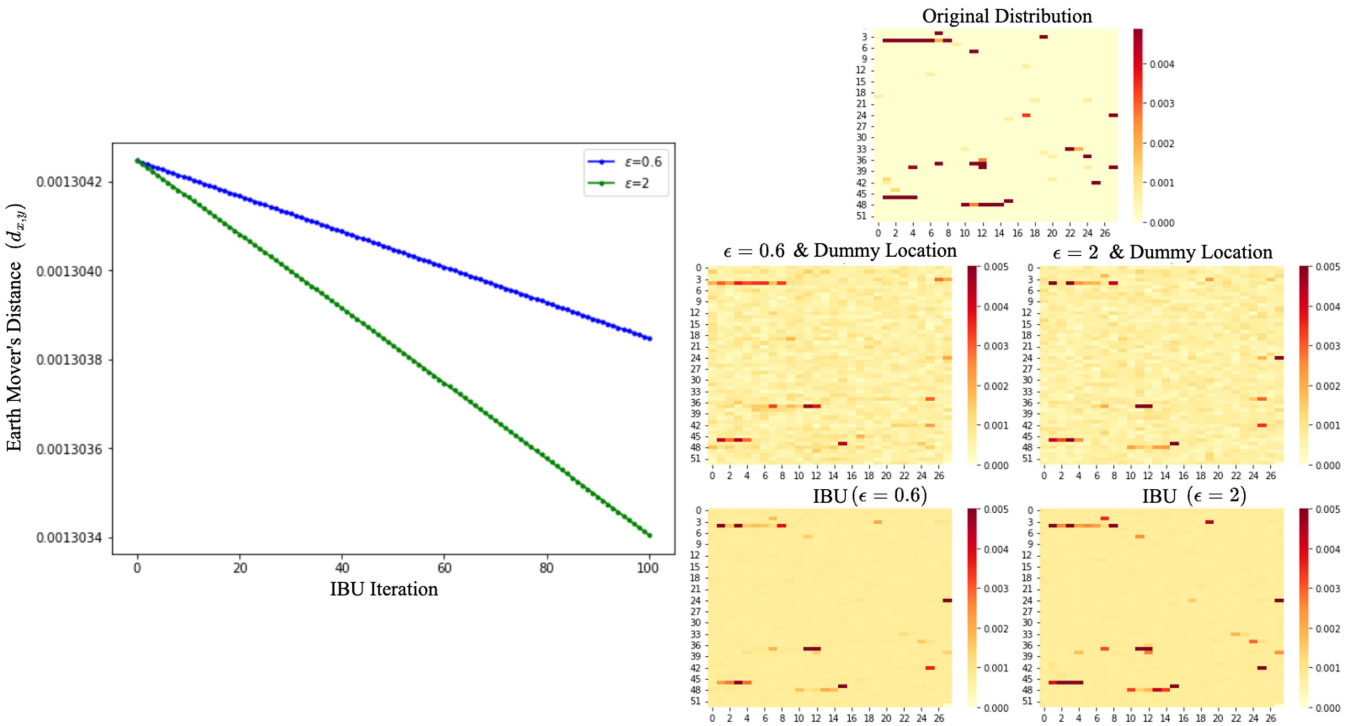


FIGURE 6. 1) Kantorovich-Wasserstein distance between the original and estimated distributions using IBU for $\epsilon = 0.6$ and $\epsilon = 2$ noisy distributions. 2) Estimations of the original distribution using IBU for the $\epsilon = 0.6$ and $\epsilon = 2$ noisy distributions.

VIII. CONCLUSION

This paper studied a fundamental problem of the risk of privacy violation for EVs dynamically querying for CS along their journeys. The setting of the problem has not been addressed in the literature, and some of the related techniques along the lines of privacy-preserving vehicle routing cannot be adapted directly into the practical model considered in this work. To address this, we theorised the notion of AGeoI allowing us to attain GeoI in a strictly bounded space of secrets. Formally justifying its soundness and applicability by proving the compositionality theorem, we derived the appropriate privacy parameters to prove that the truncated Laplace mechanism satisfies AGeoI and used it to propose a location privacy-preserving method for EVs querying for CS. Our method protects the privacy of both the specific positions of the queries and the trace of the entire journey.

In the experiments, datasets with real vehicle traces and locations were used to demonstrate the trade-off between

privacy and utility and the impact of dummy locations on this trade-off. We used IBU for real-time estimation of the original distribution of the EVs from the reported (noisy) locations. The proposed method is distinct from current machine learning-based approaches in that it considers real-time changes in the number of location-based queries. Thus, our method can capture the effect of unprecedented traffic variations on the occupancy of the CSs. Using IBU, we are capable of predicting the likelihood of a particular station being occupied by another vehicle at the time of arrival and, hence, enable an online prediction technique to estimate the availability of CS around an EV and, in turn, allowing users to do convenient route planning. A consistent trend of a substantial majority of the EVs to have "privacy for free" was observed across all the experiments, i.e., most of the EVs suffer no loss of utility even for fairly high-level formal AGeoI. In general, we observe that the cost of privacy induced by our method is fairly low across settings, thus, ensuring privacy protection for the location of the EVs without incurring a high price to pay for that. We dissected this cost of privacy incurred by our method using Voronoi decomposition to draw insight into the privacy-utility trade-off from a foundational perspective.

ACKNOWLEDGMENT

We would like to thank Prof. Graham Cormode for the insightful comments that supported the work in this paper. Ugur Ilker Atmaca and Sayan Biswas are shared co-first authors who have worked together on this paper and contributed equally. We would also like to sincerely thank the European Research Council (ERC) Advanced Grant of Catuscia Palamidessi which supported Sayan Biswas' research visit at WMG, the University of Warwick, which enabled this successful collaboration.

REFERENCES

[1] E. Ferrero, S. Alessandrini, and A. Balanzino, "Impact of the electric vehicles on the air pollution from a highway," *Appl. Energy*, vol. 169, pp. 450–459, 2016.

[2] K. Hampshire, R. German, A. Pridmore, and J. Fons, "Electric vehicles from life cycle and circular economy perspectives," *Version*, vol. 2, pp. 1–80, 2018.

[3] R. Zhang and S. Fujimori, "The role of transport electrification in global climate change mitigation scenarios," *Environ. Res. Lett.*, vol. 15, no. 3, Feb. 2020, Art. no. 0 34019.

[4] R. Hickman and D. Banister, "Looking over the horizon: Transport and reduced CO_2 emissions in the UK by 2030," *Transport Policy*, vol. 14, no. 5, pp. 377–387, 2007.

[5] S. Kufeoglu and D. K. K. Hong, "Emissions performance of electric vehicles: A case study from the united kingdom," *Appl. Energy*, vol. 260, 2020, Art. no. 114241.

[6] E. P. S. Millen, "Transport and environment statistics 2021 annual report," *Dept. Transp.*, May 2021. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/984685/transport-and-environment-statistics-2021.pdf

[7] T. Gersdorf, P. Hertzke, P. Schaufuss, and S. Schenk, "McKinsey electric vehicle index: Europe cushions a global plunge in EV sales," McKinsey Automotive and Assembly Practice, 2020.

[8] R. Hensley, S. Knupfer, and D. Pinner, "Electrifying cars: How three industries will evolve," *McKinsey Quart.*, vol. 3, no. 2009, pp. 87–96, 2009.

[9] P. Gao, R. Hensley, and A. Zielke, "A road map to the future for the auto industry," *McKinsey Quart., Oct*, pp. 1–11, 2014.

[10] M. Lombardi, K. Panerali, S. Rousselet, and J. Scalise, "Electric vehicles for smarter cities: The future of energy and mobility," *World Econ. Forum*, 2018. [Online]. Available: http://www3.weforum.org/docs/WEF_2018_%20Electric_For_Smarter_Citiespdf

[11] A. Ostermann, Y. Fabel, K. Ouan, and H. Koo, "Forecasting charging point occupancy using supervised learning algorithms," *Energies*, vol. 15, no. 9, May 2022, Art. no. 3409.

[12] A. Sao, N. Tempelmeier, and E. Demidova, "Deep information fusion for electric vehicle charging station occupancy forecasting," in *Proc. IEEE Int. Intell. Transp. Syst. Conf.*, 2021, pp. 3328–3333.

[13] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Theory of Cryptography*, S. Halevi and T. Rabin, Eds. Berlin, Germany: Springer, 2006, pp. 265–284.

[14] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor, "Our data, ourselves: Privacy via distributed noise generation," in *Advances in Cryptology - EUROCRYPT*, S. Vaudenay, Ed. Berlin, Germany: Springer, 2006, pp. 486–503.

[15] J. C. Duchi, M. I. Jordan, and M. J. Wainwright, "Local privacy and statistical minimax rates," in *Proc. IEEE 54th Annu. Symp. Found. Comput. Sci.*, 2013, pp. 429–438.

[16] M. E. Andrès, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, "Geo-indistinguishability: Differential privacy for location-based systems," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2013, pp. 901–914.

[17] N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, "Optimal geo-indistinguishable mechanisms for location privacy," in *Proc. 21th ACM Conf. Comput. Commun. Secur.*, 2014, pp. 251–262.

[18] N. Fernandes, A. McIver, and C. Morgan, "The laplace mechanism has optimal utility for differential privacy over continuous queries," in *Proc. 36th Annu. ACM/IEEE Symp. Log. Comput. Sci.*, 2021, pp. 1–12.

[19] L. Gillam, K. Katsaros, M. Dianati, and A. Mouzakitis, "Exploring edges for connected and autonomous driving," in *Proc. IEEE Conf. Comput. Commun. Workshops*, 2018, pp. 148–153.

[20] C. Maple, M. Bradbury, A. T. Le, and K. Ghirardello, "A connected and autonomous vehicle reference architecture for attack surface analysis," *Appl. Sci.*, vol. 9, no. 23, 2019, Art. no. 5101.

[21] D. Hahn, A. Munir, and V. Behzadan, "Security and privacy issues in intelligent transportation systems: Classification and challenges," *IEEE Intell. Transp. Syst. Mag.*, vol. 13, no. 1, pp. 181–196, Spring 2021.

[22] X. Lin and X. Li, "Achieving efficient cooperative message authentication in vehicular ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 7, pp. 3339–3348, Sep. 2013.

[23] N. Kumar, R. Iqbal, S. Misra, and J. J. Rodrigues, "An intelligent approach for building a secure decentralized public key infrastructure in VANET," *J. Comput. Syst. Sci.*, vol. 81, no. 6, pp. 1042–1058, 2015.

[24] P. Zhao, G. Zhang, S. Wan, G. Liu, and T. Umer, "A survey of local differential privacy for securing Internet of Vehicles," *J. Supercomputing*, vol. 76, no. 11, pp. 8391–8412, 2020.

[25] T. Franke and J. F. Krems, "Understanding charging behaviour of electric vehicle users," *Transp. Res. Part F: Traffic Psychol. Behav.*, vol. 21, pp. 75–89, 2013.

[26] R. R. Kumar and K. Alok, "Adoption of electric vehicle: A literature review and prospects for sustainability," *J. Cleaner Prod.*, vol. 253, 2020, Art. no. 119911.

[27] Z. Tian et al., "Real-time charging station recommendation system for electric-vehicle taxis," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 11, pp. 3098–3109, Nov. 2016.

[28] W. Zhang et al., "Intelligent electric vehicle charging recommendation based on multi-agent reinforcement learning," in *Proc. Web Conf.*, 2021, pp. 1856–1867.

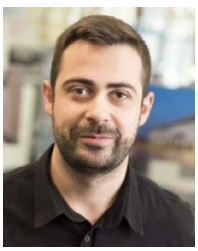
[29] R. Flocea et al., "Electric vehicle smart charging reservation algorithm," *Sensors*, vol. 22, no. 8, 2022, Art. no. 2834.

[30] G. Wang et al., "Sharedcharging: Data-driven shared charging for large-scale heterogeneous electric vehicle fleets," *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 3, no. 3, pp. 1–25, 2019.

[31] PlugShare, "Privacy policy," May 2023. [Online]. Available: https://company.plugshare.com/privacy.html

[32] ChargePoint, "Privacy and cookie policy for europe," May 2023. [Online]. Available: https://eu.chargepoint.com/privacy_policy

[33] C.-Y. Chow, M. F. Mokbel, and W. G. Aref, "Casper* query processing for location services without compromising privacy," *ACM Trans. Database Syst.*, vol. 34, no. 4, pp. 1–48, 2009.
[34] B. Niu, Q. Li, X. Zhu, G. Cao, and H. Li, "Achieving k-anonymity in privacy-aware location-based services," in *Proc. IEEE Conf. Comput. Commun.*, 2014, pp. 754–762.
[35] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Found. Trends Theor. Comput. Sci.*, vol. 9, no. 3–4, pp. 211–407, Aug. 2014.
[36] N. Guo, L. Ma, and T. Gao, "Independent mix zone for location privacy in vehicular networks," *IEEE Access*, vol. 6, pp. 16842–16850, 2018.
[37] S. Amini et al., "Cachè: Caching location-enhanced content to improve user privacy," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 14, no. 3, pp. 19–21, Dec. 2011.
[38] P. Asuquo et al., "Security and privacy in location-based services for vehicular and mobile communications: An overview, challenges, and countermeasures," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 4778–4802, Dec. 2018.
[39] L. Zhou, L. Yu, S. Du, H. Zhu, and C. Chen, "Achieving differentially private location privacy in edge-assistant connected vehicles," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4472–4481, Jun. 2019.
[40] L. Luo, Z. Han, C. Xu, and G. Zhao, "A geo-indistinguishable location privacy preservation scheme for location-based services in vehicular networks," in *Proc. Int. Conf. Algorithms Architectures Parallel Process.*, 2019, pp. 610–623.
[41] C. Qiu, A. C. Squicciarini, C. Pang, N. Wang, and B. Wu, "Location privacy protection in vehicle-based spatial crowdsourcing via geo-indistinguishability," *IEEE Trans. Mobile Comput.*, vol. 21, no. 7, pp. 2436–2450, Jul. 2022.
[42] T. Cunningham, G. Cormode, H. Ferhatosmanoglu, and D. Srivastava, "Real-world trajectory sharing with local differential privacy," *Proc. VLDB Endow.*, vol. 14, no. 11, pp. 2283–2295, Jul. 2021.
[43] D. Agrawal and C. C. Aggarwal, "On the design and quantification of privacy preserving data mining algorithms," in *Proc. 20th ACM SIGMOD-SIGACT-SIGART Symp. Princ. Database Syst.*, 2001, pp. 247–255.
[44] R. Agrawal, R. Srikant, and D. Thomas, "Privacy preserving OLAP," in *Proc. ACM SIGMOD Int. Conf. Manage. Data*, 2005, pp. 251–262.
[45] E. ElSalamouny and C. Palamidessi, "Full convergence of the iterative Bayesian update and applications to mechanisms for privacy protection," 2019, *arXiv:1909.02961*.
[46] E. ElSalamouny and C. Palamidessi, "Generalized iterative Bayesian update and applications to mechanisms for privacy protection," in *Proc. IEEE Eur. Symp. Secur. Privacy (EuroS&P)*, 2020, pp. 490–507.
[47] K. Chatzikokolakis, E. ElSalamouny, and C. Palamidessi, "Efficient utility improvement for location privacy," *Proc. Privacy Enhancing Technol.*, vol. 2017, no. 4, pp. 308–328, 2017.
[48] A. Ghosh, T. Roughgarden, and M. Sundararajan, "Universally utility-maximizing privacy mechanisms," in *Proc. 41st Annu. ACM Symp. Theory Comput.*, 2009, pp. 351–360.
[49] E. Bulut and M. C. Kisacikoglu, "Mitigating range anxiety via vehicle-to-vehicle social charging system," in *Proc. IEEE 85th Veh. Technol. Conf.*, 2017, pp. 1–5.
[50] W. Duan, J. Gu, M. Wen, G. Zhang, Y. Ji, and S. Mumtaz, "Emerging technologies for 5G-IoV networks: Applications, trends and opportunities," *IEEE Netw.*, vol. 34, no. 5, pp. 283–289, Sep./Oct. 2020.
[51] H. Ji, O. Alfarraj, and A. Tolba, "Artificial intelligence-empowered edge of vehicles: Architecture, enabling technologies, and applications," *IEEE Access*, vol. 8, pp. 61020–61034, 2020.
[52] H. Patil and V. N. Kalkhambkar, "Grid integration of electric vehicles for economic benefits: A review," *J. Modern Power Syst. Clean Energy*, vol. 9, no. 1, pp. 13–26, Jan. 2021.
[53] L. Gillam, K. Katsaros, M. Dianati, and A. Mouzakitis, "Exploring edges for connected and autonomous driving," in *Proc. IEEE Conf. Comput. Commun. Workshops*, 2018, pp. 148–153.
[54] H. Seo, K.-D. Lee, S. Yasukawa, Y. Peng, and P. Sartori, "LTE evolution for vehicle-to-everything services," *IEEE Commun. Mag.*, vol. 54, no. 6, pp. 22–28, Jun. 2016.
[55] A. Paverd, A. Martin, and I. Brown, "Modelling and automatically analysing privacy properties for honest-but-curious adversaries," Tech. Rep, 2014.
[56] "OpenStreetMap contributors," Planet dump, 2022. [Online]. Available: https://planet.osm.org
[57] A. F. D. C. U. S. Department of Energy, "Alternative fuels data center: Data downloads," May 2023. [Online]. Available: https://afdc.energy.gov/data_download/
[58] M. T. Agency, "Datasf," May 2023. [Online]. Available: https://data.sfgov.org/browse?Department-Metrics_Publishing-Department=Municipal+Transportation+Agency+
[59] M. Piorkowski, N. Sarafijanovic-Djukic, and M. Grossglauser, "CRAWDAD epfl/mobility," IEEE Dataport, 2022. [Online]. Available: https://dx.doi.org/10.15783/C7J010
[60] X. Li et al., "Perturbation-hidden: Enhancement of vehicular privacy for location-based services in internet of vehicles," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 3, pp. 2073–2086, Jul.–Sep. 2021.
[61] L. Zhang, X. Meng, K.-K. R. Choo, Y. Zhang, and F. Dai, "Privacy-preserving cloud establishment and data dissemination scheme for vehicular cloud," *IEEE Trans. Dependable Secure Comput.*, vol. 17, no. 3, pp. 634–647, May/Jun. 2020.
[62] J. Freudiger, M. Raya, M. Fèlegyházi, P. Papadimitratos, and J.-P. Hubaux, "Mix-zones for location privacy in vehicular networks," in *Proc. ACM Workshop Wireless Netw. Intell. Transp. Syst.*, 2007, pp. 1–7.
[63] H. Jiang, J. Li, P. Zhao, F. Zeng, Z. Xiao, and A. Iyengar, "Location privacy-preserving mechanisms in location-based services: A comprehensive survey," *ACM Comput. Surv.*, vol. 54, no. 1, pp. 1–36, Jan. 2021.
[64] H. Zang and J. Bolot, "Anonymization of location data does not work: A large-scale measurement study," in *Proc. 17th Annu. Int. Conf. Mobile Comput. Netw.*, 2011, pp. 145–156.
[65] C. Hecht, J. Figgener, and D. U. Sauer, "Predicting electric vehicle charging station availability using ensemble machine learning," *Energies*, vol. 14, no. 23, 2021, Art. no. 7834.
[66] A. Nait-Sidi-Moh, A. Ruzmetov, M. Bakhouya, Y. Naitmalek, and J. Gaber, "A prediction model of electric vehicle charging requests," *Procedia Comput. Sci.*, vol. 141, pp. 127–134, 2018.
[67] T.-Y. Ma and S. Faye, "Multistep electric vehicle charging station occupancy prediction using hybrid LSTM neural networks," *Energy*, vol. 244, 2022, Art. no. 123217.
[68] A. Almaghrebi, F. Aljuheshi, M. Rafaie, K. James, and M. Alahmad, "Data-driven charging demand prediction at public charging stations using supervised machine learning regression methods," *Energies*, vol. 13, no. 16, 2020, Art. no. 4231.
[69] R. Luo et al., "Deep learning approach for long-term prediction of electric vehicle (EV) charging station availability," in *Proc. IEEE Int. Intell. Transp. Syst. Conf.*, 2021, pp. 3334–3339.



UGUR ILKER ATMACA received the Ph.D. degree from the Warwick Manufacturing Group, University of Warwick, Coventry, U.K. He is currently a Postdoctoral Researcher with Alan Turing Institute, U.K. His research interests include privacy-preserving data sharing, responsible AI, threat modeling, and privacy impact assessment in an increasingly automated and interconnected world, including vehicular systems.

SAYAN BISWAS received the Master of Mathematics degree (with first-class Hons.) from the University of Bath, Bath, U.K., and the Ph.D. degree in computer science from INRIA and École Polytechnique, France. He is currently a Postdoctoral Researcher with the SaCS Lab, EPFL, Switzerland. His research interests include designing privacy-preserving techniques for analysing data and training models, with a focus on investigating and enhancing the trade-off between privacy and utility from a foundational perspective. During his Ph.D. degree, he had been a Visiting Scholar with Macquarie University, Sydney, NSW, Australia, and the University of Warwick, Coventry, England.

CARSTEN MAPLE is currently the Director of the NCSC Academic Centre of Excellence in Cyber Security Research and a Professor of cyber systems engineering with the University of Warwick, Coventry, U.K. He is also a Co-Investigator of the PETRAS National Centre of Excellence for IoT Systems Cybersecurity, where he leads on Transport and Mobility. He has an international research reputation and has authored or coauthored more than 350 peer-reviewed papers and is a coauthor of the U.K. Security Breach Investigations Report 2010, supported by the Serious Organised Crime Agency and the Police Central e-crime Unit. His research has attracted millions of pounds in funding and has been widely reported through the media. He is also a Fellow of the Alan Turing Institute, London, U.K.

CATUSCIA PALAMIDESSI (Member, IEEE) has been the Director of Research with Inria Saclay since 2002, where she leads the team COMETE. She is currently a Full Professor with the University of Genova, Genoa, Italy, during 1994–1997 and Penn State University, USA, during 1998–2002. She is coauthored more than 200 scientific publications. Her research interests include privacy, fairness, machine learning, secure information flow, formal methods, and concurrency. In 2022, she was awarded the Gran Prix of the French Academy of Science. She has been PC chair of various conferences, including Logics in Computer Science, and PC member of more than 120 international conferences. She is in the Editorial board of several journals, including the IEEE TRANSACTIONS IN DEPENDABLE AND SECURE COMPUTING, and *ACM Transactions on Privacy and Security*. She is the Chair of the ACM SIGLOG (Special Interest Group on Logic and Computation), and is in the Executive Committee of CONCUR (Concurrency Theory) and CSL (Computer Science Logic).