UTS UNIVERSITY OF TECHNOLOGY SYDNEY

Location Privacy Protection in Vehicular Networks

by Baihe Ma

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

under the supervision of Prof. Dr. Ren Ping Liu, Prof. Dr. Wei Ni, and Dr. Xu Wang

University of Technology Sydney
Faculty of Engineering and IT

September, 2023

Certificate of Authorship / Originality

I, Baihe Ma, declare that this thesis is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Electrical and Data Engineering at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signature: Production Note: Signature removed prior to publication.

Date: December 12, 2023

Thesis by Compilation Declaration				
Paper Title	**List of Authors**	**Current Status**	**Student's Contribution**	**Related Chapter**
Location Privacy Threats and Protections in Future Vehicular Networks: A Comprehensive Review	Baihe Ma Xu Wang Xiaojie Lin Yanna Jiang Caijun Sun Zhe Wang Guangsheng Yu Ying He Wei Ni Ren Ping Liu	Submitted to IEEE Communications Surveys and Tutorials. Under Second Round Review	Student's contribution is about 90%, comprising analytical work, simulations, and paper drafting. The co-authors provided feedback, comments, and technical support, which helped improve the paper.	Section I, III, IV, and V are added into: Chapter-I Introduction; Chapter-II Literature Review; Chapter-VI Conclusion.
Personalized Location Privacy With Road Network-Indistinguishability.	Baihe Ma Xu Wang Wei Ni Ren Ping Liu	Published in IEEE Transactions on Intelligent Transportation Systems 23.11 (2022): 20860-20872.	Student's contribution is about 90%, comprising analytical work, simulations, and paper drafting. The co-authors provided feedback, comments, and technical support, which helped improve the paper.	Chapter-III
New Cloaking Region Obfuscation for Road Network-Indistinguishability and Location Privacy.	Baihe Ma Xiaojin Lin Xu Wang Bin Liu Ying He Wei Ni Ren Ping Liu	Published in Proceedings of the 25th International Symposium on Research in At-tacks, Intrusions and Defenses. 2022.	Student's contribution is about 90%, comprising analytical work, simulations, and paper drafting. The co-authors provided feedback, comments, and technical support, which helped improve the paper.	Chapter-IV

Vehicle Trajectory Obfuscation and Detection.	Baihe Ma Yueyao Zhao Xu Wang Zhihong Liu Xiaojie Lin Ziwen Wang Wei Ni Ren Ping Liu	Published in book: Cybersecurity for Smart Cities. Springer, Cham, 2023. 121-134	Student's contribution is about 70%, comprising analytical work, simulations, and paper drafting. The co-authors provided feedback, comments, and technical support, which helped improve the paper.	Chapter-IV
Unlinkable Obfuscation Mechanism for Road Network-Indistinguishability and Location Privacy	Baihe Ma Xu Wang Caijun Sun Ying He Guangsheng Yu Xiaojie Lin Wei Ni Ren Ping Liu	Submitted to IEEE Transactions on Information Forensics and Security. Under Second Round Review	Student's contribution is about 90%, comprising analytical work, simulations, and paper drafting. The co-authors provided feedback, comments, and technical support, which helped improve the paper.	Chapter-V

	Name	Signature	Date
Student	Baihe Ma	Production Note: Signature removed prior to publication.	11.09.2023
Co-authors	Ren Ping Liu	Production Note: Signature removed prior to publication.	11.09.2023
	Wei Ni	Production Note: Signature removed prior to publication.	11.09.2023
	Xu Wang	Production Note: Signature removed prior to publication.	08.09.2023
	Ying He	Production Note: Signature removed prior to publication.	08/09/2023
	Zhihong Liu	Production Note: Signature removed prior to publication.	2023.9.8
	Xiaojie Lin	Production Note: Signature removed prior to publication.	2023.9.8
	Yanna Jiang	Production Note: Signature removed prior to publication.	2023.9.8
	Guangsheng Yu	Production Note: Signature removed prior to publication.	08.09.2023
	Yueyao Zhao	Production Note: Signature removed prior to publication.	2023/9/8
	Ziwen Wang	Production Note: Signature removed prior to publication.	2023.9.8
	Zhe Wang	Production Note: Signature removed prior to publication.	2023.9.8
	Bin Liu	Production Note: Signature removed prior to publication.	2023.9.8
	Caijun Sun	Production Note: Signature removed prior to publication.	2023.09.08

Abstract

Location privacy is of utmost importance in vehicular networks, where drivers' trajectories and personal information can be exposed, posing threats to drivers' safety and personal security. The proliferation of Location-Based Services (LBS) has led to a rapid increase in location data, thereby amplifying the risk to location privacy. In road networks, vehicles share location data with other vehicles and LBS through the Internet of Vehicles (IoV), making the need for effective location obfuscation techniques crucial.

Existing obfuscation mechanisms primarily focus on Two-Dimensional (2D) planar areas and overlook the unique features of road networks, often resulting in impractical outcomes such as off-road locations. Fake trajectories created by adversaries and malicious drivers can significantly compromise the utility of location data in IoV and degrade the quality of LBS. Therefore, it is essential to detect illegal trajectories to ensure the utility of location data in IoV. Some existing methods try to overcome these limitations by using pseudonyms and obfuscation, but the additive nature of differential privacy has been overlooked.

In this thesis, we propose a comprehensive differential privacy framework for protecting location privacy in vehicular networks by considering the correlation between location data and driving statuses. We first propose a personalized obfuscation mechanism that dynamically and adaptively protects the location privacy of drivers in road networks. We also define a new notion of Road Network-Indistinguishability (RN-I) to evaluate obfuscation-based mechanisms in road networks and propose a Personalized Location Privacy-Preserving (PLPP) mechanism that achieves RN-I for a single vehicle. Using the proposed RN-I, we then leverage differential privacy and propose a Cloaking Region Obfuscation (CRO) mechanism that safeguards the location privacy of multiple vehicles in road networks. To address the limitation that differential privacy makes the detection of illegal trajectories challenging, we propose a comprehensive framework for protecting location privacy in IoV by detecting illegal trajectories while preserving data utility. Finally, we introduce a new notion of Trajectory-Indistinguishability (T-I) by combining pseudonym swapping and RN-I to measure the indistinguishability of vehicles in road networks and design a Joint Trajectory Obfuscation and Pseudonym Swapping (JTOPS) mechanism that achieves T-I.

Experiments upon real-world datasets confirm the location and identity privacy-preserving capability, data utility, and efficiency of the proposed mechanisms.

Acknowledgements

This article marks the end of my doctoral career. At the same time, my days as a student will come to a close. For over twenty years, I have worked hard, struggled, been disheartened, and lost. I have always wondered: What does learning mean to me? Is it for dreams, for the future, or for the progress of humanity? Perhaps it is none of these, or perhaps it is all of them. Maybe it is to see a cloud floating by the window or to hear the noise of the night wind. Perhaps it is to realize the dream of "becoming a scientist" in elementary school. I do not know. The reasons for starting are no longer so important as I have seen the fantastic scenery.

First and foremost, I have to thank my research supervisors, Prof. Ren Ping Liu, Prof. Wei Ni and Dr. Xu Wang. Without their assistance and dedicated involvement in every step throughout the process, the publications would have never been accomplished. I would like to thank you very much for your support and understanding.

In the past three years, Prof. Liu, who is my major supervisor, has done a lot for me. I always bother him to sign very urgent papers for me. Constantly worrying about me is probably a big reason for his hair loss. I hope my graduation will minimize the damage done to his hair. Prof. Ni is the hardest-working, most enthusiastic professor I have ever met. Without him, I would never have known I could publish such papers with such quality. What I remember most about him is his hard-working: Once, when he returned from a serious illness, he said in a meeting with us, "The day you work hard is when you are in your twenties or thirties. Look at me. I am only in my forties, and my health is failing." While it is important to commit to research and love your work, I hope you will care more for your health. Is there a possibility that the story would have developed differently if you had enough sleep and were healthy enough? An academician once told me, "If you are in good health, you will be a titanic figure by the time all your peers die off." My colleague and I cannot agree more. Dr. Wang played an important role in my Ph.D. studies. He always worked day and night to help me revise my papers and polish my ideas. He also provided me with all kinds of valuable exercise opportunities. Perhaps I am an important factor for him to be single so far as worrying about me every day could waste a lot of time for his socializing. I hope that he can have more free time after my graduation, more to contact girls, and find a preferred partner as soon as possible. You know, only by trying can you have a chance, and only by taking action can you gain something. Or, we can use what you once said when you suggested Yanna and I submit our work to a top journal, "Maybe the reviewer...". You know what I mean.

I would like to thank my colleagues and co-authors. Dr. Ying and Dr. Saber helped me a lot in my research. As my seniors, they contributed valuable suggestions to my papers and corrected my mistakes. Without them, I cannot gain so much knowledge of other fields, e.g., machine learning and communication. Xiaojie Lin is the first colleague I know at UTS. We have completed many papers in the past years. Although she is stressed daily, she still works hard

and is optimistic. Please don't give up, Xiaojie. When you approach each day with optimism, life could be worse! Zixu is a good colleague who works hard and is lucky. I can feel how happy he is after marriage from his abs that have turned into a full set of abs. Yanna, who is shy and introverted, has worked with me for one year. I hope you know that you are talented. Please be confident. Bring out that same aura you had when you were joking around with me, and you'll be great. Hao Li and Haiyu Deng are also excellent. It is a pity that we did not work together, but it was still good to communicate with you. Please take care of your hair. I have noticed you guys are losing your hair a little bit. There are many friends and co-authors I want to say thanks to: Suirui Zhu, Caijun Sun, Suying Zhao, Ziwen Wang, Zhe Wang, Kaichao Shi, Yueyao Zhao, Xiaoya Guo, Prof. Zhihong Liu, and Prof. Yong Zeng. It is great to work with you. By the way, Yueyao has a boyfriend today, congratulations.

Last but not least, I would like to thank my family. Thanks to my parents (Jinghui Ma and Yanqing Qi) and grandparents (Yumei Wang and Jiang Ma) for supporting my decision. I can imagine how much you missed me in the past years because I miss you so much. Thanks to Shiyu Tan and Geng Cao for having my back. We have known each other for twenty-three years. It is great to have you guys in my life. My wife Dian Zhuang is the biggest support of my Ph.D. studies, whom I can talk and discuss everything with. Having my wife is my greatest blessing. Without her, I wouldn't have lasted today.

The flowers are blooming outside the window. It's beautiful.

Baihe Ma
January 11, 2024
Sydney, Australia

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Being an integral part of mobile systems, e.g., the fifth- and sixth-generation (5/6G) systems, vehicular networks are required to provide credible information for driving assistance [1]. Over the past decades, multifarious applications have been developed for vehicular networks to offer various Location-Based Services (LBSs). LBSs can be classified by functions into navigation, weather, venue finders, social media, and crowd-sensing [2]. By reporting location data to the LBSs, the drivers can search their destinations, check traffic conditions, and view the weather [3].

The development of LBS leads to rapidly increasing vehicular applications related to location data [4]. This poses threats to the location privacy of drivers, as LBS servers can learn the geographical locations of drivers [5]. Adversaries can collude with untrusted LBS servers to infer drivers' personal information from the location data [6]. The adversaries may also compromise trusted LBS servers or eavesdrop on the communications between the drivers and LBS servers for the location data [7], [8]. Compared with general mobile LBS users, vehicles can be relatively easily tracked because they only travel on roads and follow traffic rules [9], [10].

1.1 Location Data and Location Privacy Concerns

With increasingly popular communication sensing techniques, including cameras and vehicle-to-vehicle communications, a lot of location data will be generated in vehicular networks, in real-time and on a large scale. Location data shared in vehicular networks have characteristics such as massive, highly correlated, dynamic, and unequal importance [11]. Location data is important to help improve the safety and efficiency of vehicular networks, including vehicle infrastructure and pedestrians.

• *Massive Data:* There are enormous amounts of location information when the drivers apply LBSs [12].

• *High Correlation:* Location data is correlated, which can disclose other information when exposed.

• *Dynamic Topology:* Location data frequently changes over time.

• *Uneven Significance:* Different location information has various significance to various drivers.

Driver's home and workplace addresses tend to be more important than his shopping locations.

• *Driver Safety:* One of the most important purposes of location data transmitted in vehicular networks is safe driving [13]. The data transmission in vehicular networks should primarily satisfy the driving requirements [14]. For example, the transmission should ensure a high data utility of location data after being protected for location-sensitive LBSs, e.g., navigation.

Sharing the location data in the LBSs of vehicular networks raises privacy issues [15]. The LBSs in vehicular networks ask the drivers to expose their locations for working as planned, but untrusted LBSs and potential adversaries can lead to the leakage of the location data [16]. The adversaries can collude with untrusted LBS servers, attack trusted LBS servers, and eavesdrop on communications channels to gain the shared location data [15]. This poses the location data at risk of being obtained by adversaries. By analyzing the shared location data, the adversaries can threaten the drivers' privacy as follows.

• *Exposure of Driver's Private Information:* The adversaries can infer private information by analyzing the location data transmitted in vehicular networks, such as home addresses, religions, political parties, jobs, and work addresses [17], [18]. For example, a driver's religion can be estimated if the driver drives to religious buildings periodically. The inferred private information can be utilized to build models of targets for realizing attacks [19]. The adversaries can attack a specific driver, threaten a group of targets, or even sabotage a whole system [20].

• *Physical Attack:* The disclosure of the location privacy can lead to dangerous physical attacks, such as stalking, mugging or burglary [21]. For example, attackers can jam the traffic and induce the driver's trajectory[1].

• *Exposure of Other Vehicles' Private Information:* The drivers' trajectories can also be utilized to disclose the privacy of other drivers and predict other drivers' mobility [22]. For example, the adversaries can analyze the driver's encounter history to obtain the trajectories of other drivers.

The advanced techniques in vehicular networks introduce new threats to the driver's location privacy. With the development of vehicular networks, adversaries have a high probability of launching attacks across multiple vehicular network layers. Location privacy protection would be increasing difficult with the existing Location Privacy-Preserving Mechanisms (LPPMs) when facing cross-layer attacks. For example, the Pegasus (spyware) developed in Israel can be utilized to obtain the driver's historical encounter messages [23]. By analyzing the historical encounter messages and the eavesdropped messages, the adversaries can minimize the estimation error to achieve high-precision localization for the target driver. The development of communication capabilities allows the vehicles to share data frequently with other entities in future vehicular networks, leading to the current LPPMs, which focus on a single layer, not being able to protect location privacy when the adversaries launch attacks across multiple layers. The existing LPPMs should be improved to meet the privacy requirements in complex composite scenarios.

1.2 Location Privacy Requirement

Researchers have different definitions of location privacy [24]–[26]. In this thesis, location privacy can be defined as a branch of information privacy that deals with the location data to assist the drivers in deciding on when, what, and how to share such data [27]. The data

[1] `https://simonweckert.com/googlemapshacks.html`

shared in the vehicular networks should not threaten the driver's security [28]. The privacy requirements in vehicular networks can be classified as follows [29], [30].

• *Confidentiality:* The communication data should be well protected when a vehicle applies LBSs. The data cannot be exposed to unauthenticated parties [31].

• *Anonymity:* By analyzing the data from the same identity, the adversary can launch attacks in a long-period window. Hence, the identical information of the vehicles in the released data should be protected to avoid disclosure of the actual identity [31].

• *Unlinkability:* The link between the data and the identity to keep it safe from the adversary and unauthorized third parties [32].

• *Contextual unobservability:* The released data from the same vehicle do not branch the privacy of each other [33].

1.2.1 Characteristic of Vehicular Networks

The characteristics of vehicular networks, such as unlimited transmission power, higher computational capabilities, and predictable mobility, differ from those of other scenarios [34]. The characteristics of current vehicular networks are classified into topology features, node features, and transmission features. As shown in Fig. 1.1, the followings are the unique characteristics of vehicular networks, which are important to the location privacy in the study.

Topology Feature: The topology features of vehicular networks, related to location privacy, include mobility, dynamic network topology, real-time constraints, frequent network disconnection, and volatility.

• *Mobility:* A small delay in V2X communication can lead to severe problems since vehicles in vehicular networks move at high speeds [35]. It is impossible to employ traditional LPPMs, e.g., handshake-based authentication technologies, since most of the encountered vehicles only communicate once and do not have enough time for handshake message authentication [36]. In vehicular networks, the vehicles change their points of network attachment frequently [37]. Thus, the LPPMs in vehicular networks should have mobility management methods to meet the requirements such as seamless mobility and scalability [38].

• *Dynamic Network Topology:* The topology of vehicular networks changes quickly so that the communication duration of LPPMs is limited [39]. The density of entities on the road networks is frequently changed due to the dynamic network topology of vehicular networks that communication recourse could be unevenly distributed [40]. For example, cooperation-based LPPMs (e.g., pseudonym swap) could perform well in the city scenario rather than in the suburb scenario, as the number of nodes on the road networks in a city is much more than that in the suburbs.

• *Real-time Constraint:* The data transmission in vehicular networks is time-limited [41], and LBSs in vehicular networks (e.g., accident warning information) need time-critical messages [42]. Thus, one of the critical requirements in vehicular networks is that LPPMs should allow messages to be transmitted within an acceptable period. Nevertheless, verifying a time-critical message is difficult because the authentication process can increase the time delay [43].

• *Frequent Network Disconnection:* Vehicles frequently disconnect networks due to the high-speed movement of the vehicles and the influence of the environment [44]. A great number of

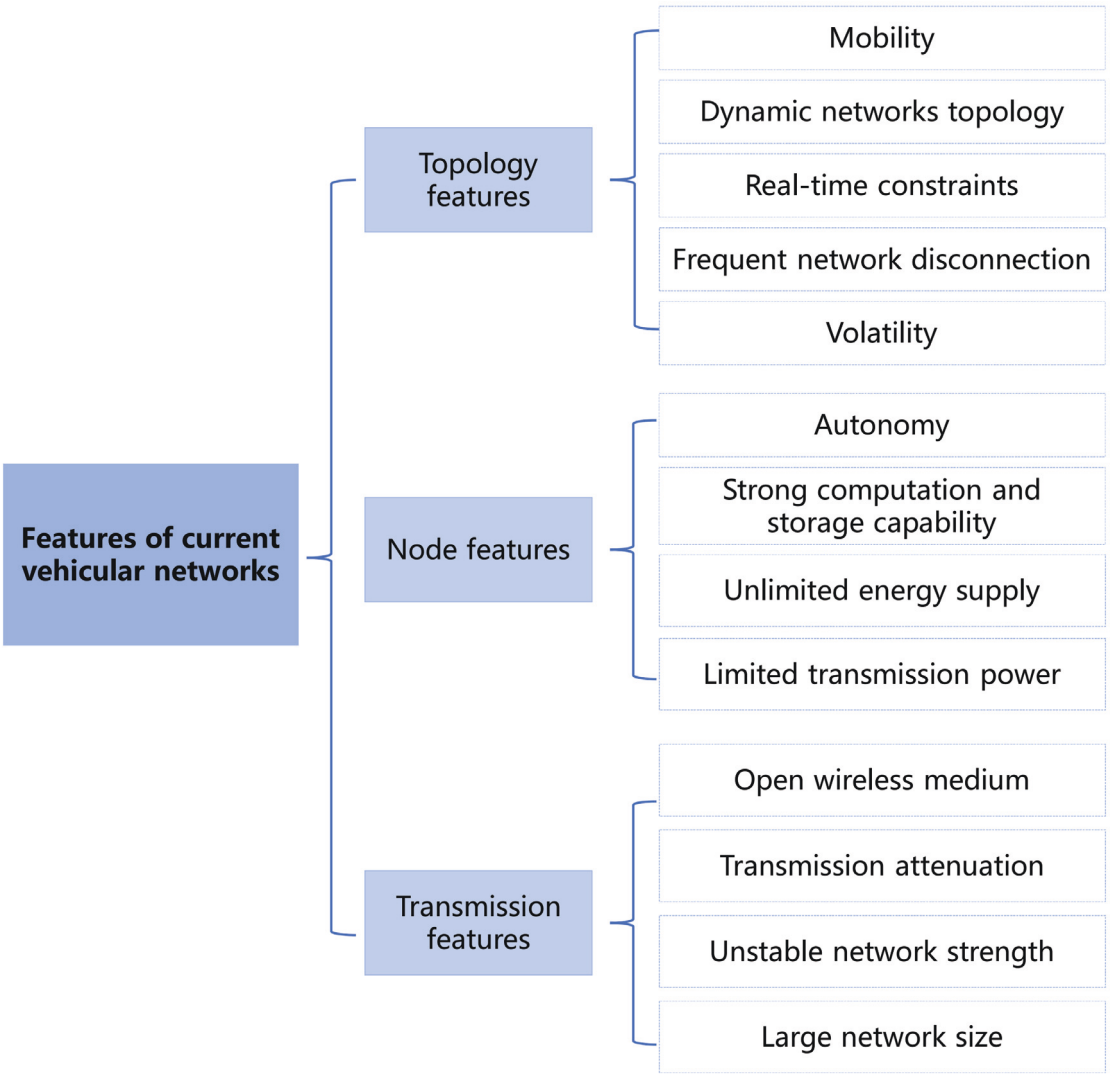


Figure 1.1: Features of vehicular networks.

vehicles in the same region can also lead to server disconnection [45]. Thus, the LPPMs should be robust to provide location privacy protection in such a scenario.

• *Connection Volatility:* The connections among vehicles are random because of the high mobility level [46]. The connectivity between two vehicles could be lost easily, and connections may remain within a specific wireless hop in a short period [47]. Vehicular networks lack a long-lived context due to the random and short connection period. Thus, it is almost impossible to utilize long-live password-based LPPMs [48].

Node Feature: The features of nodes in vehicular networks related to location privacy are autonomy, strong computation and storage capability, unlimited energy supply, and limited transmission power.

• *Autonomy:* The vehicles in vehicular networks have the authority to send, route, and receive data with limited control of centralized authority [49]. On-Board Units (OBUs) and RSUs can process the data independently [50]. Hence, decentralized LPPMs are fully considered in vehicular networks.

• *Strong Computation and Storage Capability:* The nodes in vehicular networks are required to process an extensive amount of data among vehicles and infrastructures [51]. The computation and storage capability of nodes in vehicular networks is high, so the computing resource of LPPMs could be unlimited [52].

• *Unlimited Energy Supply:* The energy consumption is unlimited in vehicular networks [53]. OBUs are supplied by vehicle battery, which satisfies the task operations [54]. Therefore, LPPMs in vehicular networks could ignore the energy limitation.

• *Limited Transmission Power:* The transmission power is limited in vehicular networks due to the characteristics of the communication protocols and the wireless access of the vehicular environment [55]. The covered area of each node in vehicular networks is limited by the transmission power. The communication region of decentralized LPPMs in vehicular networks is limited.

Transmission Features: The transmission features of vehicular networks related to location privacy are open wireless medium, transmission attenuation, unstable network strength, and large network size.

• *Open Wireless Medium:* In vehicular networks, the transmission medium is the air. Security is an important issue in vehicular networks due to the characteristics of the open wireless medium [56]. Attacks targeting open wireless mediums can also be launched to threaten location privacy in vehicular networks.

• *Transmission Attenuation:* The features of digital transmission in vehicular networks include diffraction, reflection, dispersion, refraction, and scattering [57]. These features lead to several limitations for the dedicated short-range communication in vehicular networks [58]. These features could expose the driver's location data even though the location data is well protected by LPPMs [59].

• *Unstable Network Capability:* The communication and computation capabilities of vehicle networks are unstable and affected by real-time traffic conditions [60]. For example, the network strength among vehicles can be very high in a traffic jam, as a large number of vehicles could stay in the same region and form the vehicular network. Thus, the privacy-protection capability of cooperation-based LPPMs (e.g., pseudonym swap) could be unstable.

• *Large Network Size:* Vehicular networks can cover a large area, especially downtown or on highways. However, the communication region of vehicles in vehicular networks is limited. Thus, the LPPMs should consider both the large size of vehicular networks and the small size of the communication regions.

1.2.2 Application Scenario of LPPM

The location data is protected by real-time or batch LPPMs during the collection phase and protected by offline LPPMs at the publication phase, as shown in Fig. 1.2. The data is protected by offline LPPMs if it is processed through the server to the publication stage [61], while the data is protected by real-time LPPMs before being sent to time-sensitive LBSs [62]. The batch-wise LPPMs protect location data when multiple drivers aggregate and send location data to the time-insensitive LBSs [63]. The offline LPPMs are employed when LBSs publish the trajectory data for analysis to protect the whole dataset rather than the real-time location

data [64].

1.2.3 Key Performance Index of LPPM

Three major metrics are defined by the existing works to evaluate and compare the performances of different LPPMs as follows [2], [65], [66].

• *Privacy:* The privacy metric measures the location-privacy protection capability of the LPPMs under the adversaries' attacks [15], [21], [67].

• *Data Utility:* The data utility metric is determined by the LBSs' and drivers' requirements. The privacy metric and the data utility metric are conflicting [65], [68]. For example, the data utility will achieve the peak when the location information is unprotected, and vice versa.

• *Efficiency:* The efficiency metric measures the time and storage cost of the LPPMs [69], which uses computational overhead, storage overhead, scalability, and tolerance of error as metrics [70], [71].

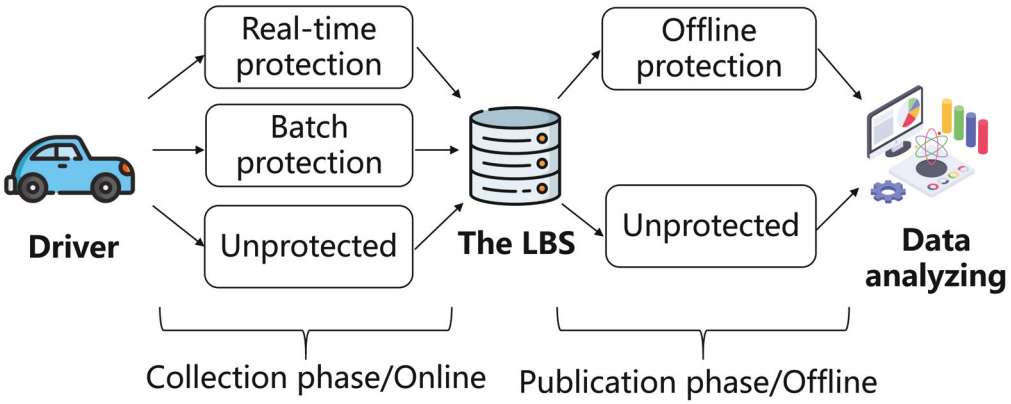


Figure 1.2: The cases of employing LPPM to protect location privacy.

1.3 Motivation

Obfuscation-based schemes can protect location privacy against untrusted LBS servers and eavesdroppers [72]–[74]. The idea is based on the local differential privacy technology where users locally perturb raw data with differential privacy schemes and employ the obfuscated data rather than the accurate data [75]. In obfuscation-based schemes, drivers generate obfuscated locations by adding controllable noise in their actual locations and report the obfuscated locations in LBS requests [69]. The obfuscated locations are selected according to the distances between the obfuscated and actual locations under the constraint of indistinguishability [74]. Although the obfuscation-based schemes compromise location accuracy, they can still be used in location-insensitive LBS, such as location-based recommendations [76].

Obfuscation-based schemes have not been rigorously developed for vehicles in road networks. In general-purpose obfuscation schemes, e.g., two-dimensional (2D) Laplace location privacy-preserving schemes, the distance between an actual location and the corresponding obfuscated location is measured using the Euclidean distance [73], [74], [77]. This, however, underestimates

the distance between two locations on road networks because the route distance is no shorter than the Euclidean distance between any two locations [78]. Another issue of general-purpose obfuscation schemes is that they may generate off-road obfuscated locations, for example, on a river [79]. Additional steps can be introduced to avoid off-road locations, which, however, increase the computational complexity [80].

The privacy-preserving levels of locations are expected to be fine-tuned across road networks to balance data utility and location-privacy protection [81]. Drivers can be sensitive to a small number of locations in road networks. If all locations are obfuscated at the highest level as the sensitive locations, the insensitive locations are over-obfuscated, and the data utility is penalized [82]. If only the sensitive locations are obfuscated, the sensitive locations are statistically different from the insensitive locations. This compromises location privacy [83].

In existing obfuscation-based mechanisms, malicious drivers breach legal drivers' profit. For example, malicious drivers can occupy more benefits than they deserve by deliberately modifying their trajectory data in Taxi service [84], [85]. Malicious drivers also use location privacy-preserving schemes to protect their location data. By analyzing location data which concludes illegal location data, smart city applications cannot provide an acceptable QoS of LBS. Therefore, LBS should detect illegal location data to ensure high QoS. If the malicious drivers employ location privacy-preserving schemes (e.g., obfuscation schemes) as the legal drivers, detecting illegal data becomes difficult.

The temporal information has not been well considered in road networks by the existing obfuscation mechanisms [86]. If an adversary can track the identities (ID) of the vehicles, it can reduce the perturbation noise by linking the communications of a specific vehicle ID over a long-period, such as long-observation attack [87]. As a result, the existing obfuscation-based mechanisms can hardly protect location privacy on a long-term basis [88].

1.4 Research Contributions and Overview

This thesis aims to bridge the gaps between the existing LPPMs by focusing on real-world road networks.

Chapter 2 provides a comprehensive review of existing LPPMs relevant to this thesis. We first illustrate the existing localization techniques which introduce challenges to the existing LPPMs. We then review the existing LPPMs by comparing their strengths and limitations. Finally, we discuss the LPPMs to balance location privacy and data utility in vehicular networks.

In Chapter 3, we define a new measure of location indistinguishability for road networks by applying the concept of differential privacy. The indistinguishability is measured only on road networks (as opposed to continuous 2D space). Route distances (i.e., distances along roads) are used to derive the differential privacy upper bound. Off-road locations are precluded from obfuscation operations. We propose a new dual-obfuscation algorithm that first probabilistically obfuscates an actual location to a connection and then obfuscates the location into a road interval between the actual location and the obfuscated connection. By carefully designing the probabilities, we prove that the dual-obfuscation design satisfies the new differential privacy-based definition of road network-indistinguishability. Considering the non-uniform location privacy requirements of a driver at different locations, we apply the nearest neighbor interpolation to specify the privacy budgets for all locations based on the sensitive locations of the driver. The vehicle locations can be obfuscated consistently without exposing the sensitive

locations, while improving the utility of the location data of less sensitive locations. We conduct comprehensive experiments on two real-world trajectory datasets and compare the proposed PLPP scheme with the existing 2D Laplace location privacy-preserving schemes [73], [77], [89]. Experimental results show that the PLPP scheme outperforms the 2D Laplace schemes in terms of data utility, privacy-preserving level, and efficiency.

In Chapter 4, we research the mechanisms that improve data utility and privacy protection in road networks. We employ differential privacy in the proposed personalized obfuscation scheme to protect drivers' location privacy adaptively and to provide high Quality of Services (QoS) of LBS in road networks. We propose a Convolutional Neural Network (CNN) based detection mechanism to detect illegal trajectories without requiring the drivers' actual location. The proposed scheme has high accuracy in detecting illegal trajectories even if the drivers protect actual location data with various noise sizes. We further protect location privacy by cooperating with multiple vehicles in road networks. Firstly, we develop differential privacy in road networks and extend the proposed RN-Indistinguishability that quantifies road network location privacy with unique road network features. By using the extended RN-Indistinguishability, we design the CRO mechanism that employs the route distances to quantify the indistinguishability of locations on roads. We prove that the CRO mechanism satisfies the RN-Indistinguishability. The CRO mechanism can be extended with general road network features without breaching differential privacy. Then, we use pseudonym swapping to incorporate multiple vehicles in the time domain and the spatial domain. We demonstrate that pseudonym swaps can be considered a differential privacy process for any two identities sharing the same pseudonym-swapping candidate set. We analytically prove that jointly using pseudonym swapping and obfuscation in vehicular networks can achieve higher privacy protection than separately using them.

In Chapter 5, we introduce a new unified privacy-preserving measure, i.e., T-I, treating trajectory obfuscation and pseudonym-based mechanisms as a holistic process. T-I extends the applicability of ε-differential privacy (DP) by quantifying the distinguishability between vehicles based on their historical information in the time domain and current information in the spatial domain. Building upon T-I, we propose a novel Joint Trajectory Obfuscation and Pseudonym Swapping (JTOPS) mechanism joining pseudonym swapping and trajectory obfuscation with a crafted criterion. The proposed mechanism is proven to combine two differential privacy processes, without introducing the additivity composition theorem of ε-DP. The proposed pseudonym swapping does not require vehicles to disclose their private data, as the required parameters are computed locally and then transmitted to the coordinator. Thus, the knowledge of the coordinator, which assists in pseudonym swapping, is limited, ensuring the proposed pseudonym-swapping process can effectively resist collusion attacks.

Chapter 6 concludes the thesis and outlines our future works, while Chapter 7 lists the publications of the author.

Chapter 2

Literature Review

In this chapter, the existing tracking techniques in the different scenarios are first summarized from the aspect of location privacy, as shown in Table 2.1. Then, we illustrate the localization requirements, list the general adversary models, and provide possible privacy threats in current vehicular networks. By highlighting the threats introduced by localization techniques, the adversaries that have been well-considered are presented. Then, we assess the existing LPPMs in three categories and discuss the limitations of each category, as shown in Table 2.2. The existing LPPMs are divided into user-side, server-side, and user-server-interface LPPMs according to the phase that the LPPMs are allocated, by considering the threats of localization techniques [90]. The user-side LPPMs allow the drivers to process data before sending it to the LBSs, while the server-side LPPMs process the aggregated location data in dataset. The user-server-interface LPPMs use trusted third parties and secure communication to ensure that the location data is secure in transmission. We summarize the limitations of LPPMs from the aspect of localization and communication requirements, and we review the method to balance location privacy and data utility from the aspects of theory and practice. The theoretical methods are classified into the blockchain, adaptive parameters, hybrid, encryption, elements simplification, and virtual nodes, where we provide the existing works to illustrate the method in detail. The real-world in-vehicle tracking techniques, i.e., COVID-19 tracking applications, are used as examples to discuss the balance between location privacy and quality of service in practice.

2.1 Localization and Tracking Techniques in Vehicular Networks

2.1.1 Localization and Tracking Techniques

Over the past decades, tracking techniques have been presented for high-precision LBSs and anti-theft systems [127]. However, the tracking techniques can be used by the adversaries, which branches the drivers' location privacy. Multiple LPPMs are presented to prevent the adversaries from obtaining the drivers' privacy.

The existing tracking techniques can be classified as sensing infrastructure-based, optical vision-based, vehicle driving log-based, cellular radio-based, and upper-layer message-based. The adversaries with upper-layer message-based tracking techniques are well considered in the existing LPPMs, while the adversaries with vehicle driving log-based tracking techniques are

Table 2.1: Existing localization techniques and their challenges.

Category	**Technique**	**Requirement**	**Description**	**Challenge**
Sensing infrastructure-based tracking	Magnetic sensors [91]–[95]	Obtain magnetic induction signal generated by the vehicles	Calculate three-dimensional location and two-dimensional orientation information of the vehicles	Malicious sensors are difficult to be detected; The adversaries can collude with sensors; The adversaries can hijack legal sensors; No existing work can defend the adversaries with sensors
	Inductive loop detectors [96], [97]	The inductive loop generated by the vehicles	Can be deployed at intersections on road networks; Can monitor information like speed, volume, and size	
	Beacon message-based sensors [98]–[100]	Beacon messages like AVI tags, RFID tags, GPS, and MAC	Extract private information from beacon messages	
Optical vision-based tracking	ML and AI [101]–[103]	Use ML and AI to analyze the geometrical information of the captured camera frames	Observe the physical features of the vehicles; Image identification background	It is impossible to hide the physical features of the vehicles; Image identification has a high accuracy; The adversaries can directly observe the vehicles in the real world; No existing LPPM can defend the adversaries in such a scenario

Vehicle driving log-based tracking	Reverse engineering [104]–[106]	GPS and CAN data	Read CAN data from OBD and use reverse engineering technique to calculate location data	Multiple OBD readers can be combined for localization
	In-vehicle communication [107]–[115]	In-vehicle communication system; Drivers' mobiles	Hijack in-vehicle communication system to obtain location data; Hijack mobile phones to read location data	Few existing LPPMs focus on hardware layer location privacy; An in-vehicle communication system is difficult to be considered in LPPMs for LBSs; The mobile devices are almost impossible to be isolated in the vehicles
Cellular radio-based tracking	Antenna localization [116]–[119]	Base stations Antenna signal	Use the direction of the antenna signal for localization	Antennas are equipped in the vehicles for communication; Signals can be detected in the physical world; Based station is necessary for communication; Localization with multiple based stations is very easy; No existing works can defend cellular radio-based localization

Upper-layer message-based tracking	Estimation actual information [120]–[126]	Messages in vehicular networks	The adversaries can launch multiple attacks to obtain messages that are transmitted in vehicular networks, with which it can infer the drivers' actual locations;	Various attacks can be utilized. The adversaries can launch attacks that cross multiple layers

overlooked by the existing studies. Location privacy under sensing infrastructure-based, optical vision-based, and cellular radio-based tracking can almost not be protected.

Sensing Infrastructure-based Tracking

Sensing infrastructure-based tracking techniques detect vehicle trajectories by using equipment such as inductive loop, infrared, ultrasonic, microwave, magnetic and piezoelectric sensors [128]. The driving features captured by the sensors can be utilized for tracking.

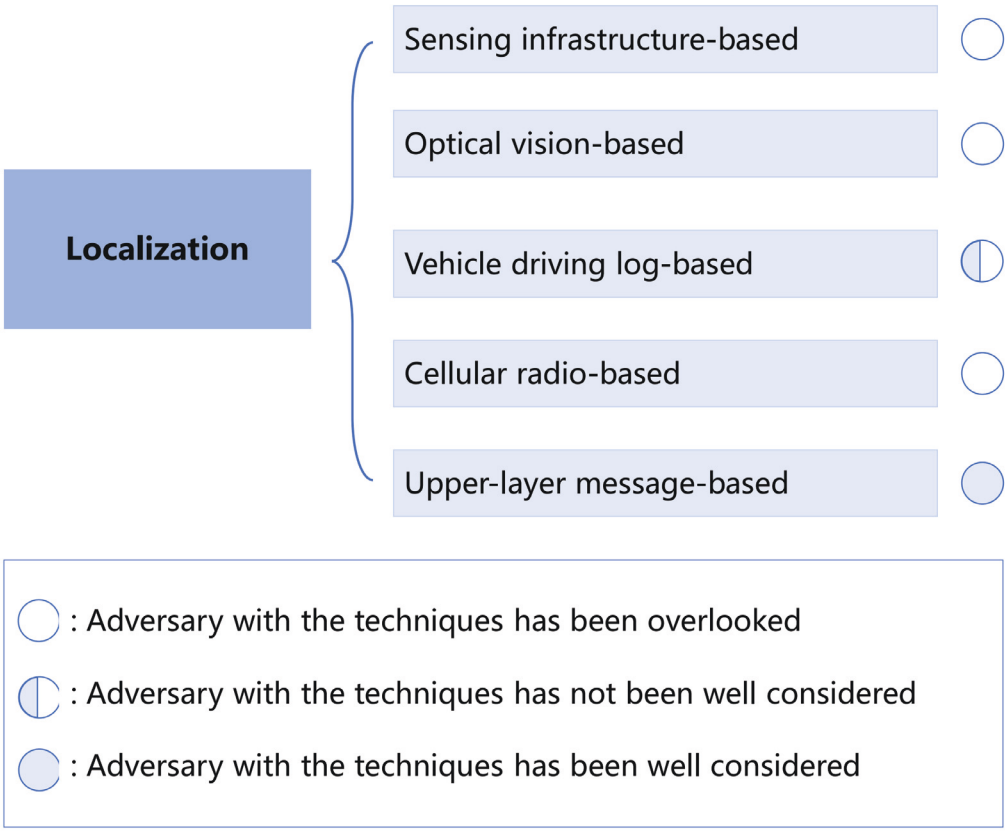


Figure 2.1: Localization techniques.

• *Magnetic Sensor:* Vehicles disrupt the Earth's magnetic field that generates magnetic induction signals [129]. The magnetic sensors can expose the 3D location and 2D orientation information of the vehicles [91]. The magnetic sensors can capture the generated signals to

analyze the travel time and vehicles' identifications [92]. The existing works [92]–[95] modeled the magnetic field perturbations caused by the vehicles and extracted the magnetic waveforms of the vehicles. With the models and the waveforms, the techniques can track the vehicles and infer the driving status of the vehicles.

• *Inductive Loop Detector:* The inductive loop detectors are widely employed to obtain traffic data, which can be deployed at intersections on road networks to monitor traffic [130]. The inductive loop detection can provide vehicle information, such as speed, volume, and size [129]. The existing works [96], [97] used the data obtained from inductive loop detectors to increase the tracking accuracy.

• *Beacon-message-based Sensor:* The Beacon-message-based tracking techniques utilize messages like Automatic Vehicle Identification (AVI) tags and Radio Frequency IDentification (RFID) tags, Global Positioning Systems (GPS), and Medium Access Control (MAC) [131]–[134]. The existing works [98]–[100] tracked vehicles with multiple beacon-message-based sensors because a single sensor is susceptible to the vehicles' driving status and environment.

No existing LPPMs can protect drivers' location privacy under the sensing of adversaries who utilize multiple sensors. The existing studies focused on protecting location data captured by the legal sensors through techniques such as blockchain, obfuscation, and anonymity.

Optical Vision-based Tracking

The optical vision-based tracking estimates vehicles' locations by analyzing the geometrical information of the captured camera frames (i.e., image pixels) [101]. The optical vision-based tracking can provide a high-precision estimation and robustness system in low-textured and low-visibility environments [135].

The existing works utilized ML and AI techniques to extract appearance features, such as color and texture, and combined the extracted features with other semantic information for tracking [129]. Study in [102] developed a deep learning framework based on the cross-frame keypoint-based detection network and spatial motion information-guided tracking network. The developed framework considers the vehicles' driving status and appearances extracted from satellite videos. Yang et al. [103] improved a collaborative sensing system that integrates a customized metric-learning vision-based vehicle re-identification method to extract vehicle features. The system combines vehicle appearance with traffic network connection as features, which achieves a high accuracy of tracking.

To the best of our knowledge, no existing LPPM can prevent the adversaries from tracking vehicles based on the appearance features.

Vehicle-Sensor-based Tracking

On-Board Diagnostic (OBD) readers can be utilized to obtain the data of Electronic Control Units (ECU) in the Controller Area Network (CAN) [136]. The easiest way to locate vehicles is by reading data from the GPS devices, but vehicle manufacturers have divided CAN into sub-networks to harden the vehicles against single sensor tracking [104], [105]. Nevertheless, multiple OBD readers (e.g., inertial, heading, pressure, and speed sensors) can be combined to realize localization [106], [137].

It has been demonstrated in the literature that in-vehicle communication systems such as CAN,

entertainment applications, and drivers' mobile phones can be used to locate vehicles [107]–[114]. For example, the Tegaron from Daimler Chrysler has navigation infrastructures and can exchange data with Tegaron's control centre [107]. In this way, vehicle localization can be achieved either by the in-vehicle communication systems from the third party or by the mobile communication linkage [108]. The Tire Pressure Monitor System (TPMS), as part of the in-vehicle wireless networks, is another potential attack surface to track vehicle locations [109]. The identifiers of the TPMS sensors are unique, and the protocol is vulnerable to reverse engineering, so the Road-Side Unite (RSU) can launch a passive tracking system to capture TPMS packets and map a route with the unique sensor identifiers [110], [111].

Mobile phones with GPS play an important role in the side channel attack to infer the driving trajectory. The adversaries can use the accelerometer data from the mobile [112] or the angle matching of the route with the mobile magnetometer sensor data [113] to track a car. Guha et al. [114] leveraged the data from the accelerometer and gyroscope sensors to detect vehicular movements, stops and turns to match the driving path. Besides the existing vehicle on-board hardware, vehicular services and applications can also provide the information to track vehicle trajectory in the side-channel attack. The in-vehicle auxiliary applications are required to support the autonomous driving functions by providing services in the passengers' services, vehicles' services, intelligent communication options, and intelligent resource allocation [115]. However, these in-vehicle auxiliary applications will generate more data embedding the vehicle state and location, which increases the privacy breach risk.

To the best of our knowledge, there are no universal solutions defending against location privacy attacks in the hardware-based physical layer. The existing in-vehicle intrusion detection systems, such as gateway-, ECU-, and CAN-based [138], focus on detecting malicious data transmitted in-vehicle and can hardly prevent the above passive side-channel attacks on location privacy.

Cellular Radio-based Tracking

Antenna systems are designed to track and steer signals from vehicles [139]. For tracking systems and localization purposes, it is most desirable to combine a small compact omnidirectional sensor array using beamforming techniques [140]. However, the omnidirectional antenna is susceptible to signal loss in long-distance communication [139]. In the case of long-distance communication, the existing studies utilized a directional high gain antenna that focuses signal energy in a particular direction [116]. Burghal et al. [117] realized a relative vehicular localization using the channel state information from multiple-antenna transceivers. The authors used feed-forward neural networks to reduce the number of trainable parameters. The mmWave communication can also contribute to vehicle localization, as electrically steerable directivity of phased arrays offers a direction finding [118]. Several direction findings of base stations can be merged in a cross-bearing manner to localize the vehicles [119].

To the best of our knowledge, no existing studies have considered the adversaries with cellular radio-based localization.

Upper-Layer Message-based Tracking

The upper-layer message-based tracking techniques allow the adversaries to obtain vehicles' location data by eavesdropping on the V2X communications or attacking the entities. In this subsection, we discuss upper-layer message-based tracking techniques by analyzing the

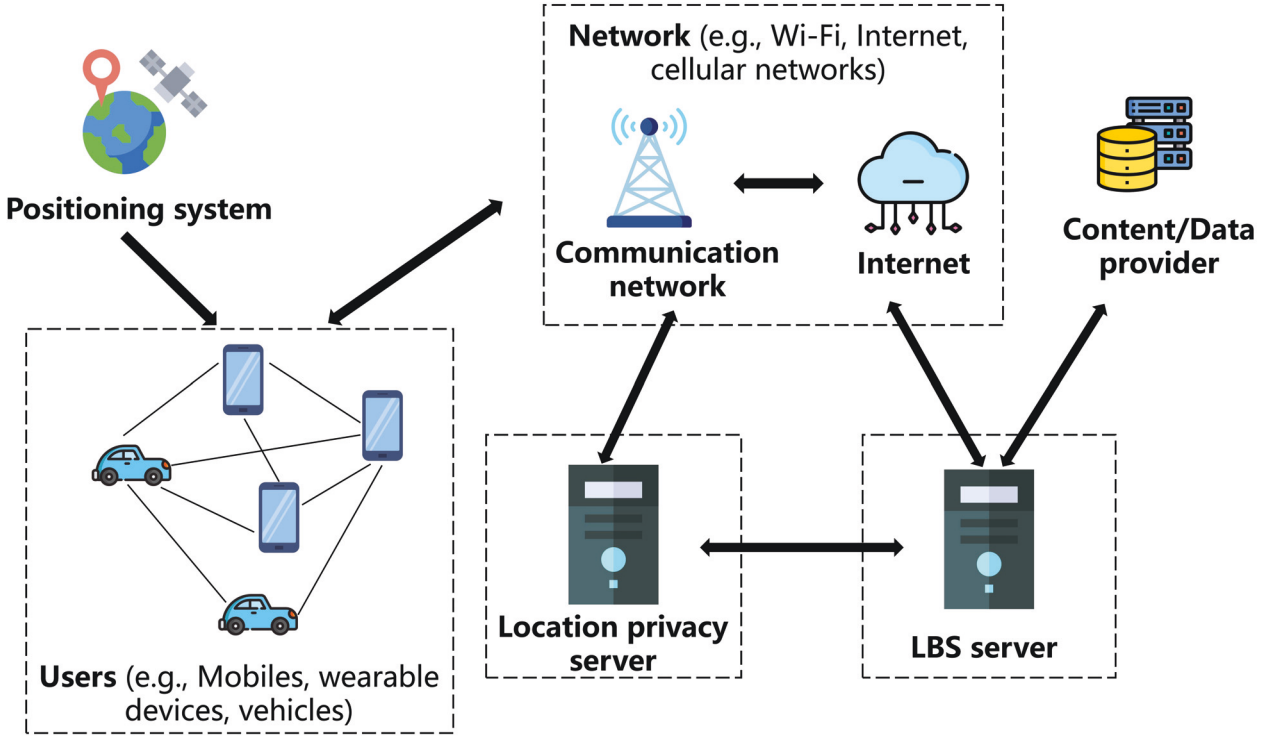


Figure 2.2: System model of LBS.

adversaries mentioned in most existing works, as shown in Fig. 2.1. The adversaries are classified based on the behaviors and the size of the monitor region.

There are three components in an LBS system: drivers, servers, and networks, as shown in Fig. 2.2. Each part of the LBS system can be threatened by the adversaries. For example, eavesdropping on the network can disclose the messages transmitted between two nodes.

Based on the behaviors, the adversaries can be classified into active attackers and passive attackers [120]. The active adversaries focus on disrupting network communication, falsifying communication data, and inserting fake messages by capturing or cloning the legal drivers [141]. Compared with the active adversaries, the passive adversaries aim to monitor and analyze the data traffic to discover drivers' positions by eavesdropping [121]. The passive adversaries do not disrupt or interfere with communication directly [122]. The majority of existing LPPMs focus more on the passive adversaries than the active ones.

According to the size of the region that the adversaries can monitor, the passive adversaries can be classified as Global Passive Adversaries (GPA) and Local Passive Adversaries (LPA) [123], as shown in Fig. 2.3. GPA can eavesdrop on data transmitted in networks with full knowledge of the road by monitoring data, obtaining legitimate authorities or hacking into applications, and can gather information over a long time, like hours, months, or even years [7]. The GPA model is considered as strongest adversary in many privacy methods [124]–[126]. However, the GPA is a theoretical scenario because the GPA model needs to place a large amount of equipment with a prohibitive cost [121]. Compared with the GPA, the LPA eavesdrop on a relatively small communication range with limited equipment.

Fig. 2.4 shows an overview of the four main steps of location attacks. The adversaries obtain the location information of the drivers by collecting, eavesdropping, and compromising the vehicle's information. By analyzing the obtained information, the adversaries can gain useful knowledge,

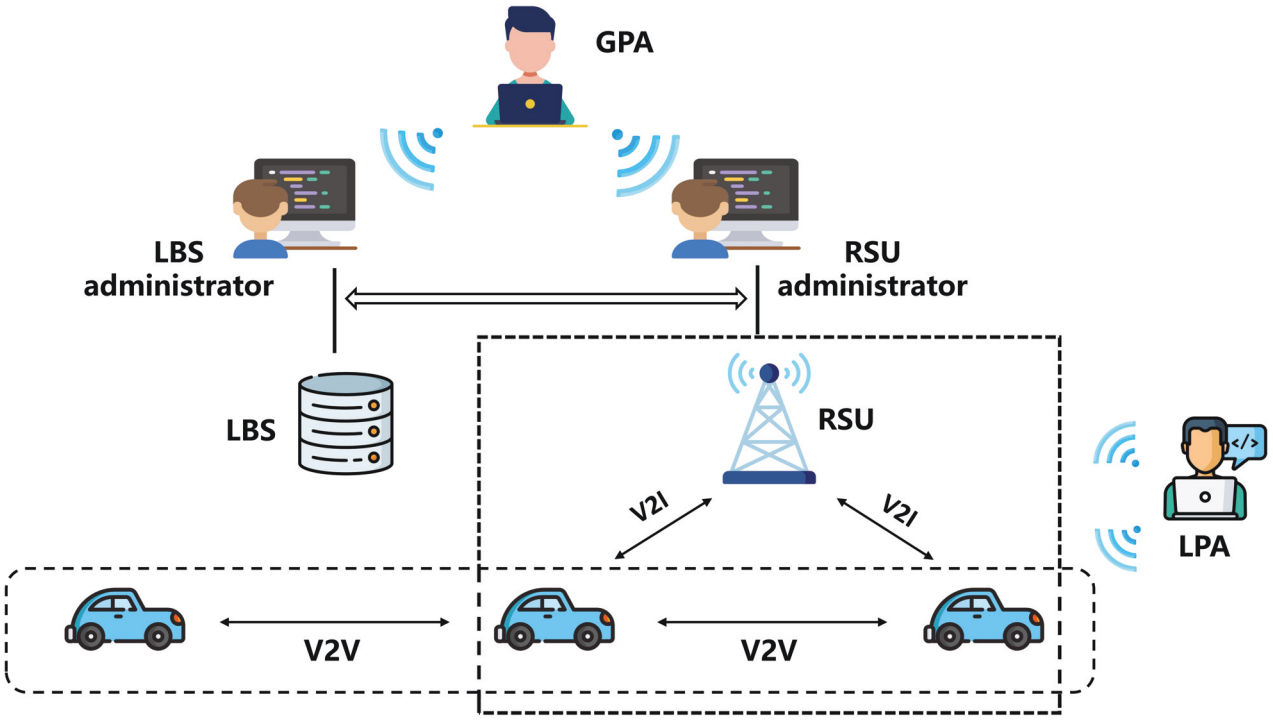


Figure 2.3: GPA and LPA for vehicle location privacy.

with which the adversaries can use methods, e.g., context linking, probability theory, machine learning, and (fake) peer user to infer the location information of the driver. And then, the identities, discretized trajectory points, and continuous trajectories would be exposed to the adversaries.

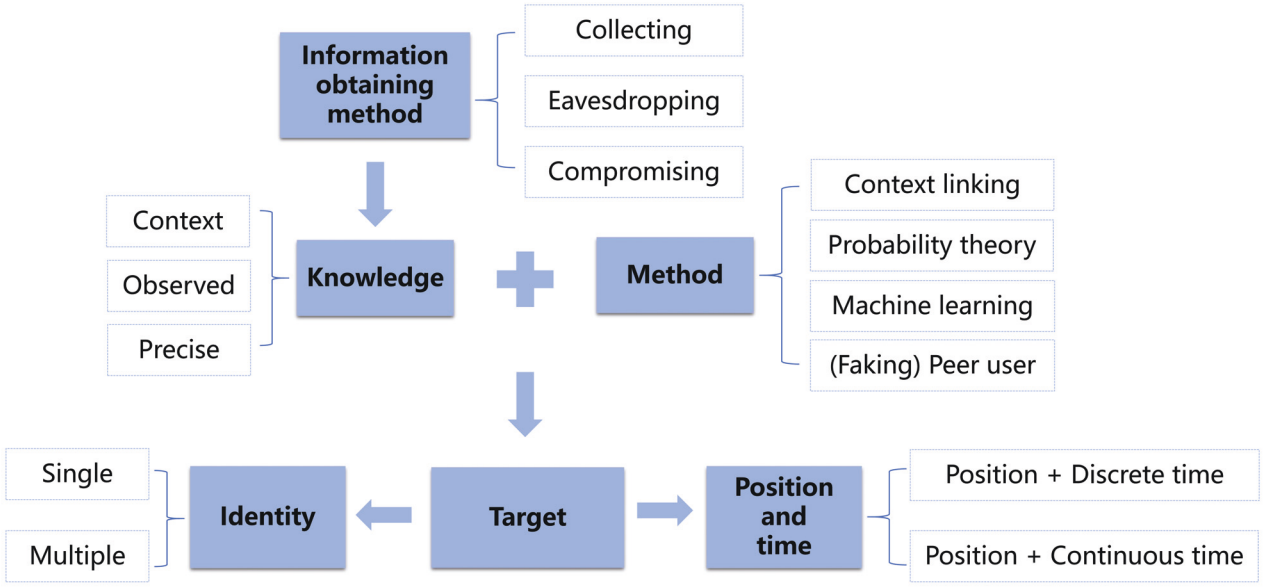


Figure 2.4: Overview of the location attacks and adversaries.

2.1.2 Location Privacy Threat in Vehicular Networks

The existing vehicular networks are under various location privacy attacks. One of the simplest methods to track a vehicle is illegally setting a Global Navigation Satellite System (GNSS) on it [142]. The adversaries can obtain the drivers' real-time locations by this method. However, this method is infeasible because the adversaries have to set GNSS equipment for each target vehicle.

Passive and Active Attack

The location privacy attacks can be viewed as a typical Multiple Target Tracking (MTT) issue, which assumes a set of noisy measurements or observations detected periodically by a sensor [143]. The adversaries obtain the best estimation of the driver's state and association probability through various location privacy attacks. The location privacy attacks can be classified into passive attacks and active attacks [141], as shown in Fig. 2.5.

The adversaries with passive attacks aim to monitor and analyze the data traffic to discover the drivers' positions by eavesdropping that do not disrupt or interfere with communication directly [121]. The passive attacks in vehicular networks can be classified as follows [141].

• *Wireless Eavesdropping Attack:* The adversaries can easily eavesdrop on vehicular networks due to the open wireless medium. Through eavesdropping, the adversaries can obtain data transmitted in networks for future analysis [144].

• *Tracing Back Attack:* The adversaries can employ a triangulation algorithm to locate the drivers' positions with at least two separate antennas through tracing back attacks [145].

• *Traffic Analysis Attack:* The core nodes in vehicular networks process more traffic flow than ordinary nodes [146]. For example, sources and destinations transmit data at a higher rate speed than ordinary nodes. Thus, the adversaries can discover the targets' locations by analyzing the traffic flow.

• *Packet Analysis Attack:* The adversaries can extract encrypted information from the captured packet, e.g., location information and identification (ID) information [147]. The adversaries can infer the sender's trajectory by analyzing packets' timestamps if the packets have the same ID in two different locations.

• *Back-rolling Attack:* The adversaries can utilize the back-rolling attack if they have larger storage than legitimate nodes [141]. The adversaries can record the drivers' historical and new positions to find out the location information of targets.

The active attackers can interfere with vehicular networks (e.g., compromising and cloning legal nodes). The active attacks can be classified as follows [141].

• *Node Compromised Attack:* The active adversaries can capture packets from several legitimate nodes to estimate drivers' locations [148]. The adversaries can also infer the network topology and drivers' locations by interfering with networks through cloning nodes [149].

• *Routing Blocking Attack:* The routing-blocking attacks track back the target drivers hop by hop to obtain the drivers' data. The adversaries can trigger a fixed routing block and monitor the data through routing blocking attacks [150].

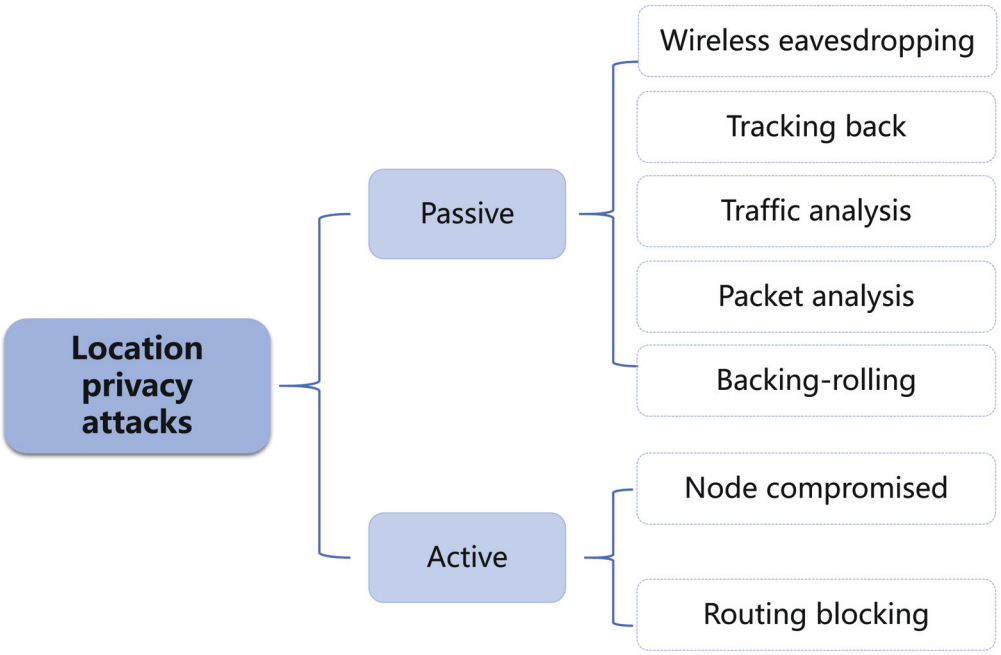


Figure 2.5: Major privacy attacks.

Attack based on Physical Characteristic

The physical characteristics can disclose the drivers' location information. As shown in Fig. 2.6[1], the adversaries can track the drivers' locations by observing vehicles' shapes, license plates, and other characteristics. The tracking technologies (e.g., Automatic Number Plate Recognition (ANPR)) make it possible to track the vehicles with their physical characteristics [151]. The adversaries can estimate the vehicle's trajectory when the vehicle is identified at several different locations. The adversaries' estimation error decreases with the number of exposed locations increases. The obtained data can be utilized for later tracking attacks even if a vehicle is identified at one location [152].

Attack on Inside-vehicle Message

Inside-vehicle communication messages indicate the drivers' location data since the messages represent vehicle speed and steering angle [153]. The adversaries can obtain the data in Electronic Control Units (ECU) by illegal On-Board Diagnostic (OBD) reader on the OBD port or through vulnerabilities [154]. Inside sensors can transmit data in wireless networks [142]. For example, the tyre pressure monitoring system can broadcast unencrypted data within 40 meters. Remote keyless entry technology sends unique identifiers to unlock vehicles through short-range broadcasts. The unique identifiers can be used to track the vehicles [155].

Attack on V2X Message

V2X communications in vehicular networks are at the risk of disclosing location privacy, as shown in Fig. 2.7. Cooperative Awareness Messages (CAM) and Decentralized Environmental Notification Messages (DENM) are periodically broadcasted by intelligent transport system stations [156]. CAM and DENM messages consist of the vehicles' status information, such as

[1]The vehicle in the figure is owned by the authors.

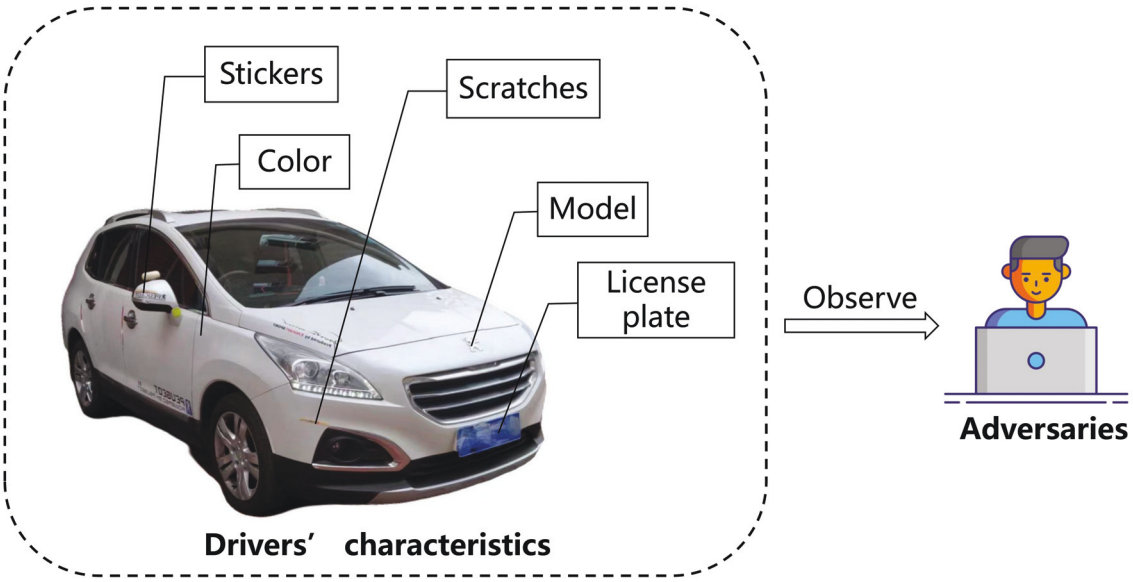


Figure 2.6: The identifiable characteristics of the vehicles for tracking.

time, location, and speed. CAM and DENM are necessary for the security services in vehicular networks, e.g., road condition warning and hazardous location warning [157]. CAM and DENM, which are unencrypted to decrease time delay, contain the drivers' digital signatures for message verification [158]. The digital signatures of CAM and DENM disclose the vehicles' accurate location data through high-time resolution [142]. The medium access control (MAC) layer protocols can disclose drivers' locations, such as the time-division multiple access (TDMA) [159]. TDMA MAC slot is broadcasted as the vehicles' identifications, which makes it easy for the adversaries to track the vehicles' trajectories by eavesdropping on the wireless channels [160].

The additional devices and peripherals inside the vehicles can also disclose the drivers' location privacy, as shown in Fig. 2.7. The adversaries or telecommunication companies can recognize the drivers' regions when they use their mobile phones within the region of a signal tower [161]. And then, the adversaries can achieve high-precision localization by using multiple signal towers. Another unique identification that can be utilized for tracking is International Mobile Subscriber Identity (IMSI) [162]. IMSI can be captured by multiple IMSI catchers, which lead to vehicle tracking if the drivers take their mobile phones into their vehicles [163]. Bluetooth sensors in mobile phones also periodically broadcast advertisement packets containing the devices' MAC addresses [164]. The adversaries can calculate the drivers' trajectories if they can record the Bluetooth MAC addresses of the drivers [165].

The adversaries can track vehicles without the content of the transmitted data. As shown in Fig. 2.5 and Fig. 2.6, the signal direction context is sufficient for the adversaries to realize trace-back attacks within the wireless sensor networks [166]. In vehicular networks, the adversaries can also infer the position changes and velocity changes with received CAM.

Security Attacks in Vehicular Networks

In the intricate ecosystem of vehicular networks, security attacks manifest through a spectrum of tactics, ranging from data breaches to unauthorized remote control of vehicles and severely compromising network integrity and vehicular operation. The attacks can lead to serious privacy and security issues, including unauthorized access and malicious activities aimed at both

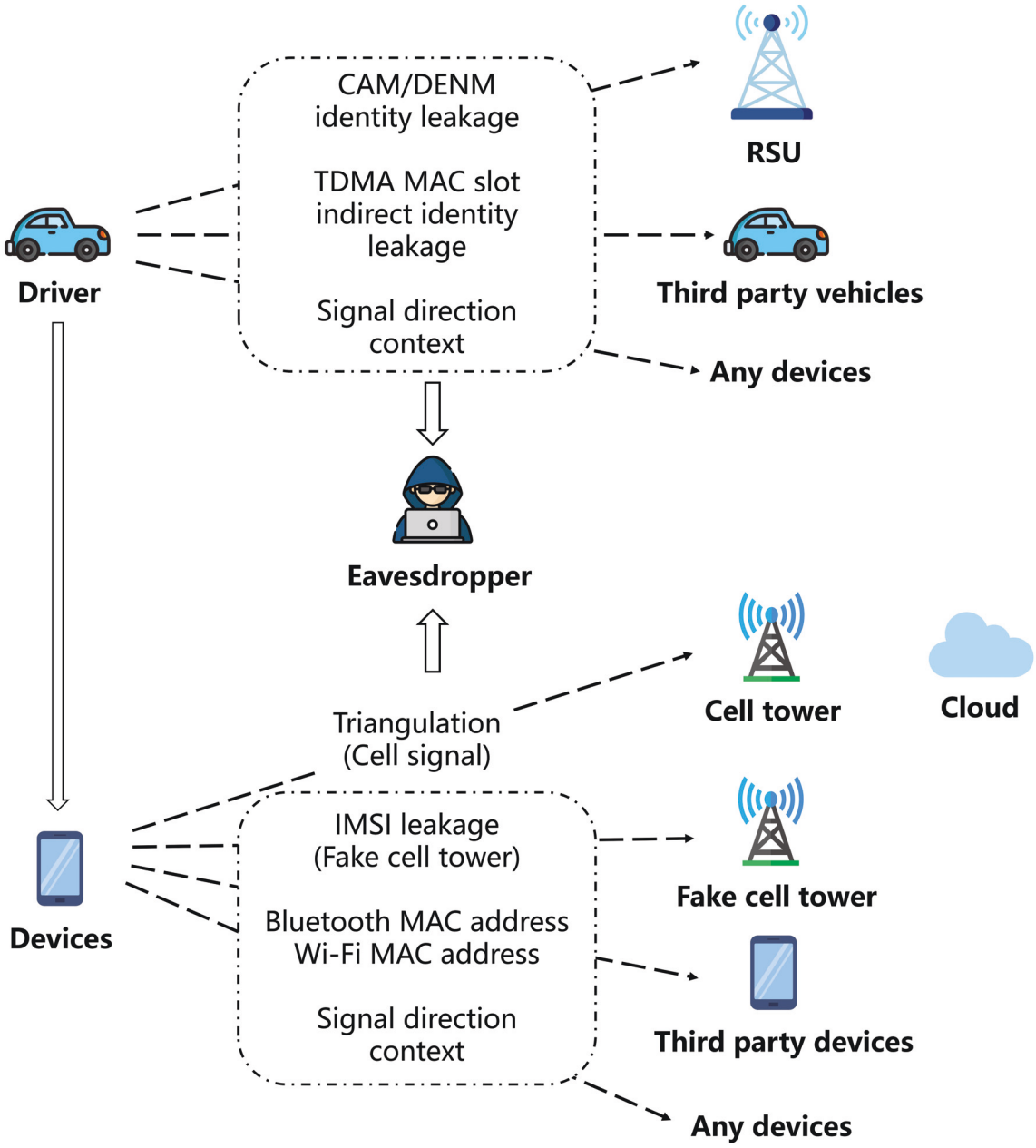


Figure 2.7: V2X and in-vehicle communication threats.

data and physical control of vehicles. These attacks can override or bypass a driver's decisions, leading to potential adverse remote control or manipulation of vehicle functions. In platoon scenarios, cybersecurity breaches could result in a loss of formation control, unauthorized lane changes, or platoon disbandment—each with potentially catastrophic outcomes. Real-world instances, like certain Sydney roads granting green lights to approaching trucks to maintain traffic fluidity, highlight the potential impact of such attacks on city logistics [2]. Beyond general data privacy, individual vehicle tracking poses direct threats to driver safety and privacy, with identity-based and message replay attacks being prime examples of such vulnerabilities in vehicular communications.

Security measures adopted in vehicular networks strive to safeguard data against unauthorized

[2] https://www.governmentnews.com.au/new-technology-sees-trucks-talk-to-traffic-lights/

access, manipulation, or disclosure, fostering trust within the network [167]. The array of security attacks, detailed in [168], [169] are multifaceted.

• Identity-based Attacks: Adversaries may spoof vehicle identities, employing tactics like Sybil attacks to create counterfeit vehicles with fabricated identities [170] or use GPS spoofing to generate deceptive location data [171]. These masquerading attacks conceal the attacker's true identity, allowing them to send fraudulent data and disrupt network integrity. Additionally, Man-in-the-Middle (MITM) attacks enable adversaries to eavesdrop and respond maliciously in communications, effectively impersonating legitimate vehicles [172].

• Message Modification-based Attacks: Vehicular networks can be compromised by attacks that alter message integrity, such as black hole or gray hole attacks [173], where adversaries drop or unpredictably manipulate data routing [174]. Illusion attacks are particularly insidious, sending erroneous traffic data to legal vehicles and potentially causing accidents by influencing vehicular behavior [175].

• Message Replay-based Attacks: Adversaries may undermine network topology correctness through replay attacks, injecting outdated but valid messages to deceive vehicles [176]. Wormhole attacks involve adversaries creating a low-latency data transmission tunnel, misrepresented as a high-quality route, to intercept and control vehicular data [177].

• Network Availability-based Attacks: In this kind of attack, an adversary aims to exhaust the resources of vehicular networks that legal vehicles cannot access services [178]. Denial-of-Service (DoS) attacks flood the network with excessive data. In contrast, Distributed Denial-of-Service (DDoS) attacks and jamming attacks, which are two widely observed variants of DoS, can be executed from multiple locations or through noise interference, respectively [178]. The former allows the adversary to launch the attack from different locations, while the latter relies on noise generated by the adversary.

2.2 LPPM in Vehicular Networks

We assess the existing LPPMs in three categories and discuss the limitations of each category, as shown in Table 2.2. The LPPMs are classified into user-side, server-side, and user-server-interface LPPMs. User-side LPPMs process location data during the collection phase, while server-side LPPMs protect location privacy during the publication phase. The user-server-interface LPPMs use trusted third parties and secure communications to realize location privacy protection. We summarize the limitations of LPPMs from the aspect of localization and communication requirements, and we review the method to balance location privacy and data utility from the aspects of theory and practice. The theoretical methods are classified into the blockchain, adaptive parameters, hybrid, encryption, elements simplification, and virtual nodes, where we provide the existing works to illustrate the method in detail. The real-world in-vehicle tracking techniques, i.e., COVID-19 tracking applications, are used as examples to discuss the balance between location privacy and quality of service in practice.

2.2.1 User-Side LPPMs

The user-side LPPMs aim to protect location privacy on the driver side in the collection phase. The popular user-side LPPMs include pass-and-run, certificate, secure computation, and data perturbation.

Pass-and-Run

Pass-and-run aims to transmit the information through other vehicles instead of sending data directly to LBSs [179]. The pass-and-run is first proposed in [180]. The authors treat vehicular networks as delay-tolerant networks. A vehicle can decide whether to pass the message to other vehicles or submit the message to LBSs. Lu et al. [181] propose a lightweight pass-and-run method where the location data is perturbed according to the mobility of vehicles and the delay of transmission. The vehicles have two strategies in the method, i.e., a greedy strategy and a random strategy, which are selected based on the drivers' requirements. The authors utilize an asymmetric encryption algorithm to protect the passed message. Nevertheless, the transmission delay of the method is high.

Table 2.2: An overview of the existing LPPMs.

Category	**Techniques**	**Description**	**Benefits**	**Limitations**
User-side LPPMs	Pass-and-Run	Transmit location data through other vehicles	Transmission delay breaks the linkage of location information	High communication complexity
	Certificates for Privacy	Authentication	Provide both location privacy and authentication	High computational and storage consumption
	Secure Computation	Operate encrypted data directly	Operation is flexible and does not reveal private data	High computational and transmission delay
	Data perturbation	Apply LBSs with fake location data rather than the actual version	Considers the prior knowledge of the adversaries and does not need a trusted third party	Low data utility

Certificates for Privacy

Certificates, also known as privacy-preserving attribute-based credentials, are cryptographic mechanisms [182] that allow the drivers to obtain certified credentials for their attributes from trusted issuers, only reveal required information satisfying the requested LBSs' predicates [183]. The design of certificates relies on the use of malleable signatures [184], as follows.

• *Attribute-based Signature:* Shahandashti et al. [185] introduce the concept of attribute-based

Server-side LPPMs	Statistical Disclosure Control	Anonymize or obfuscates data in the dataset	Keep the general statistical features of the dataset	Need a trusted third party
	Homomorphic Encryption	Operate encrypted data directly on the server side	The operation could be flexible and does not need to decrypt data	High computational consumption and transmission delay
	Private Information Retrieval	Hide the requested items	The drivers can apply for LBSs without disclosing their requirements	High computational consumption
	Searchable Encryption	Hide plaintext keywords in searching.	Can be combined with other methods	Low accuracy
User-server-interface LPPMs	Secure Communic-ation	Use traditional protocols or end-to-end encrypted services	Techniques have been developed in past decades	There are some limitations of the traditional protocols
	Trusted Third Party	Introduce trusted third parties to assist communication	High efficiency	It is an ideal environment

signature, which allows the drivers to sign messages with fine-grained control over identifying information [186]. The attribute-based signature is calculated based on the attribute value, which can be a binary-bit string [187], [188], or relies on a particular data structure [189]. Maji et al. [186] and El Kaafarani et al. [190] introduce and formalize the anonymity of the attribute-based signature that the signature should not reveal the driver's identity or the used attributes. The authors point out that the adversaries can track the drivers' locations by identifying their actual attributes. Considering the same privacy issues, Kaaniche et al. [191] improve the anonymous attribute-based signature by using a concrete mathematical construction based on standard assumptions and the random oracle model. The improved model significantly improves the location privacy and identity privacy of the attribute-based signature.

• *Group Signature:* Group signature allows a driver of a group to anonymously sign messages on behalf of the group [192]. Then, any verified vehicle can confirm that the signature is generated by a legal group member without requiring to identify the signer. Zheng et al. [193] improve a linkable group signature that achieves anonymity, auditing, and tracing functions for the communication sender. The improved framework increases communication efficiency and security by employing cryptography modules of blind signature, public-key encryption, trapdoor indicative commitment, and signature of knowledge. Hakeem et al. [194] employ a short-size signature to broadcast authentication information over multiple zones. The authors decrease the signature generation time and verification time by utilizing bilinear pair cryptography. Wu et al. [195] point out that bilinear pair cryptography is complex for OBUs and RSUs. The authors use an elliptic curve cryptosystem in the authentication process that decreases the computational complexity of group signatures. Mundhe et al. [196] develop a ring signature that provides unconditional privacy to the drivers by transmitting the message through the verified legal vehicles. The authors also utilize the pseudonym to increase the privacy-preserving capability of the ring signature. Mundhe et al. [197] enable RSUs to participate in the signature generation process to decrease the authentication delay of the ring-signature-based LPPMs. Mei et al. [198] adopt the full aggregation certificateless signature technology to reduce the bandwidth resources overhead in the transmission of the certificate.

• *Sanitizable Signature:* Ateniese et al. [199] are the first to develop Sanitizable signatures in 2005. In the sanitizable signature, authorized semi-trusted censors can modify part of a signed message in a limited and controlled fashion without interacting with the original signer. Pamies et al. [200] combine the log anonymity and sterilizable signature to protect the sensitive data, e.g., identity and location. The authors hide authenticated identification when transferring the data streams from the local node to remote storage servers.

• *Blind Signature:* For the blind signature, the content of a message is disguised before it is signed, which means the signer and message author are different parties [201]. Sun et al. [202] improve a fog-computing-based crowdsensing architecture where identity and location privacy are protected by a partially blind signature authentication. The authors use zero-knowledge verification to improve the security of the blind signature.

In summary, the limitation of the signature-based methods is that certificate management leads to high computational and storage consumption.

Secure Computation Mechanism

Secure computation techniques protect a driver's location privacy by processing the location data, which was first introduced and formalized in 1982 based on the millionaire problem [203].

Zhuo et al. [5] improve a multi-key secure outsourced computation scheme that does not need the interaction between the LBS servers and drivers. The improved scheme avoids duplicating and useless encrypted LBS messages before verifying identities, which ensures the quality of LBSs.

Data Perturbation

The obfuscation-based LPPMs allow the drivers to perturb their location before sending the location information to LBS servers [15]. The data perturbation can be utilized on either the user side or the server side. In user-side data perturbation, drivers do not need to trust any external entity by sending obfuscated location data to LBS servers [204]. The data perturbations can ensure personalized location privacy, which offers location privacy protection with acceptable data utility [15]. The user-side data perturbation is summarized as follows.

• *Dummy-based Method:* Dummy-based methods do not need a trusted third party or key sharing step [70]. Niu et al. [205] improve a dummy-location selection algorithm to achieve k-anonymity for the drivers. The algorithm selects dummy locations according to the entropy of anonymity. The limitation of the dummy-based methods is the insecurity of submitting spatio-temporal correlation in sequential LBS requests. Liu et al. [206] point out that time reachability, direction similarity, and the degree should be a concern when filtering the dummies to solve this limitation. After the filtering strategy, the candidate dummies can provide high privacy protection for the driver's location information. The storage cost of the filtering strategy is acceptable, but the computation delay could be high.

• *Local Differential Privacy (LDP):* Differential privacy can bind the knowledge obtained by the adversaries but could decrease the quality of LBSs [207]. From the LBS providers' perspective, applications can provide a high quality of services if high accuracy is received, but most LBSs can accept location data that is not entirely accurate [208]. The differential privacy can be defined as follows,

Definition 1 (Differential Privacy (DP))**.** *A mechanism $\mathcal{M}$ satisfies ϵ-DP if and only if, for any pair of data x_i and x_j, we have*

$$\frac{\Pr[\mathcal{M}(x_i) \to y]}{\Pr[\mathcal{M}(x_j) \to y]} \leq e^{\epsilon}, \tag{2.1}$$

where y is the output of the mechanism.

LDP is a distributed variant of traditional differential privacy that allows the drivers to perturb their location information before sending it to servers [209]. Erlingsson et al. [210] amplify the privacy to achieve high privacy-preserving capability from local differential privacy by combining differential privacy and anonymity. The authors point out that location privacy security can be achieved without adding any significant noise if the method employs LDP on the client side and a shuffling strategy on the server side. Nevertheless, the authors ignore the spatio-temporal correlation in the road network.

• *Geo-Indistinguishability (Geo-I):* Geo-I is first proposed by Andrés [211]. Geo-I allows the drivers to enjoy ϵr-differential privacy in the given obfuscate radius r with a privacy budget ϵ, which is as given by

Definition 2 (Geo-I)**.** *A mechanism $\mathcal{M}$ satisfies ϵ-Geo-I if and only if, for any two locations x_i and x_j, the following holds*

$$\frac{\Pr[\mathcal{M}(x_i) \to y]}{\Pr[\mathcal{M}(x_j) \to y]} \leq e^{\epsilon d(x_i, x_j)}, \tag{2.2}$$

where y is the output of the mechanism and $d(x_i, x_j)$ is the distance of x_i and x_j.

Based on the Geo-I, Zhou et al. [212] design a framework to balance the utility and privacy in edge computing. The framework includes two parts: privacy-preserving location-based service usage method and privacy-based service adjustment. The authors add two-dimensional Gaussian noise to shift actual locations. Although the authors consider the trade-off between privacy and service quality, the framework can lead to a high calculation consumption. Based on the background knowledge of trusty serves, Li et al. [213] employ correlation probabilities and correlation transition probabilities to realize Geo-I. The proposed method can provide different privacy-preserving levels for various requirements. The location shift decides the level of privacy protection. The shift is invalid when the driver's location shift is shorter than the threshold. Although the method can provide a different protection level, the method ignores the road condition. If the traffic jams and the driver applies LBSs frequently, the obfuscated location will never change by which the adversaries can obtain the driving state. Li et al. [214] improve an enhanced Geo-I definition named Perturbation-Hidden to ensure perturbed locations are valid. The Perturbation-Hidden method transforms the map of road networks into a grid where the acceptable locations are used as the candidate set. Furthermore, dynamic programming is employed to determine the retrieval area to provide accurate LBSs. The authors employ the dynamic programming method to provide the drivers with the shortest radius of retrieval radius. The limitation of the Perturbation-Hidden method is that it may lead to high privacy costs in privacy-limited regions.

• *Pseudonym:* Pseudonyms are employed as temporary anonymous certificates generated and distributed by the certificate authority [215]. The pseudonym-based LPPMs aim to ensure the unlinkability between the driver's identity and pseudonyms in communication [216]. Wang et al. [217] treat pseudonyms as a long-term identifier to decrease the computation and storage consumption. The authors design a trigger of pseudonym exchange requests to assist the certificate authority in the pseudonym changing. Vehicles change their pseudonyms when meeting the trigger. The method has a limitation that a long period of existing pseudonyms may provide a longer track window for the adversaries. Pseudonym-Indistinguishability is first proposed in [218] to ensure strict unlinkability in the pseudonym swap process. The pseudonym swap process satisfies differential privacy. The adversaries cannot link the pseudonyms after swapping, even if the driving states of the two vehicles are similar. The Pseudonym-Indistinguishability method can provide a high swap complete probability with fewer pseudonyms. The limitation of the method is that the authors ignore the conflict in the pseudonym swap. There are two weaknesses of the pseudonym-based LPPMs. One weakness is that the pseudonym-based LPPMs need to manage the vehicles' pseudonyms, which leads to high computation and storage consumption [219]. Another weakness is that the pseudonym-based LPPMs cannot ensure unlinkability in the tracking attack [220].

2.2.2 Server-Side LPPMs

By using server-side LPPMs, service providers are required to perform additional processing on their clients' hosted data [221]. Service providers can achieve this by anonymizing databases,

removing identifying traces or encrypting data contents.

Statistical Disclosure Control (SDC)

SDC mechanisms are mainly used to protect data within statistical databases, which balances data utility and drivers' location privacy [222]. Generally, the output of the SDC mechanisms ensures that the databases do not reveal information related to a specific driver [223]. Database anonymity and differential privacy are the two most popular techniques in the SDC mechanism, as follows.

• *k-anonymity:* k-anonymity LPPMs can provide personalized and accurate query results without key sharing [224]. k-anonymity LPPMs protect driver's location data in k vehicles to increase the estimation error of the adversaries [225]. The k vehicles are selected according to the closest historical request probability based on the maximum entropy principle [226]. The existing k-anonymity methods can be classified into centralized k-anonymity and distributed k-anonymity [227]. The centralized k-anonymity methods need anonymous cloaking servers, while the distributed k-anonymity methods ask the participants to unite [228]. Li et al. [229] propose a distributed architecture utilizing the blockchain. The proposed architecture records the hashed safety beacon messages to ensure integrity by reducing storage consumption and processing time. The authors employ k-anonymity to gather and upload safety beacon messages of a group of vehicles. The proposed architecture does not need a trusted third party. Luo et al. [17] improve a blockchain-enabled k-anonymity method to ensure trustworthiness. The proposed method calculates the trust level of vehicles based on the historical trust data and the trust degree reflecting factors. The authors establish a blockchain structure to record the historical trust information. The authors treat RSUs as distributed k-anonymity databases that provide the recorded data.

• *Mix-zone:* The Mix-zone methods allow a certain number of drivers to change pseudonyms in specific regions [230]. The drivers orderly get into the regions and get out of the regions in a different order [231]. Mix-zone methods cannot against timing and transition attacks that the adversaries can link the driver's pseudonyms at entry and exit points based on the timing knowledge [232], [233]. The adversaries can also employ the continuous query correlation attack to link the old and new pseudonyms [234]. Some pseudonym mappings can be ruled out because of the weighting time, traffic conditions, and time constraints [67]. Amro [235] proposes a Mix-zone method that introduces fixed transceivers to decrease the influence of traffic conditions. The legitimate virtual vehicles join in the pseudonym swap process when the number of physical vehicles is lacking. The proposed method requires transceivers that act as physical vehicles. The introduced transceivers use pseudonyms to communicate with RSUs. However, the introduced transceivers bring some security and privacy issues. For example, the pseudonym swap process will be insecure if the adversaries obtain the pseudonyms of the transceivers. A group-based dynamic Mix-zone method is proposed to protect location privacy in resources limited regions [236]. The proposed method allows vehicles to transmit encrypted information in the silent region. The method is personalized that considers the expiration time of the drivers' pseudonyms.

• *Other Anonymity-base LPPM:* Meng et al. [237] propose a method to protect the driver's location privacy in navigation services. The method extends anonymous authentication that supports a request-limiting property. The method protects both location and route information of the users. The method can effectively protect identity privacy, location privacy, route privacy, unlinkability, and confidentiality. Zhu et al. [238] propose an Anonymous Smart-parking

And Payment (ASAP) method to realize smart parking navigation. The ASAP method applies servers anonymously that the drivers transmit their locations into a region with the cloaking method to hash and encrypt the location data. Locations and routes can be explored under similar continuous queries, although the actual location is cloaked. Singh et al. [239] propose a Masqueraded Probabilistic Flooding for Source-Location Privacy (MPFSLP) method to reduce communication costs. The method can ensure non-repudiation, message authentication, integrity, and non-traceability. In the MPFSLP method, the re-sending is defined to replace the forwarding process. The authors point out that the vehicles do not require a relation between the location data and the identities of other vehicles. The MPFSLP method allows each node to send packets by masquerading as probabilistic flooding instead of generating fake packets. The vehicles can re-send previous messages to complicate the identity.

• *General Differential Privacy:* Differential privacy can also be used on the server side. Soheila et al. [240] propose a Differentially Private Data Streaming (DPDS) system to aggregate the data in vehicular networks. By considering the data correlation, vehicles in the DPDS system set up several groups in which the group leaders send members' compressed information. The compressed information reduces the noise value, but its limitation is that the size of the compressed data is similar to the original data. In [241], the authors propose a protection framework to protect privacy in edge computing. They also propose a data transmission method based on the noise quadtree and the Hilbert curve. The proposed method can improve the efficiency of location data publishing. Nonetheless, they only consider the two-dimensional space, which may not satisfy the real-world environment.

• *Joint Differential Privacy:* Joint differential privacy limits the single driver's manipulation power and reduces the impact of the single driver's false report [242]. Based on the joint differential privacy, a study in [243] proposes a scheduling protocol to protect location privacy and minimize vehicle miles in the ridesharing services. The proposed method employs private dual decomposition technology and driver clustering to improve the scheduling performance. They also design the private ride assignment method and driver grouping method to improve the privacy-preserving capability.

Homomorphic Encryption

Encryption is widely used to protect the location privacy of vehicles in data dissemination and computation, but data decryption needs to be more flexible and secure [244]. Homomorphic Encryption (HE) schemes can directly analyze ciphertexts by mirroring the corresponding operations on the plaintexts, which means the decrypted values of computation results on ciphertexts of location data in the HE schemes corresponds to the values of operations on the plaintexts [245]. The most common definition of the HE is given by

Definition 3. *Let $\mathcal{P}$ and* C *be a set of plaintexts and ciphertexts, respectively. An encryption mechanism $\mathcal{M}$ satisfies homomorphic if and only if, for any given encryption key k, and any pair of data x_i and x_j, the following holds: $\forall x_i, x_j$*

$$\mathcal{M}(x_i \odot_{\mathcal{P}} x_j) \leftarrow \mathcal{M}(x_i) \odot_{\mathrm{C}} \mathcal{M}(x_j), \tag{2.3}$$

for some operators $\odot_{\mathcal{P}}$ in $\mathcal{P}$ and $\odot_{\mathrm{C}}$ in C*, where $\rightarrow$ means that the left-hand side can be directly computed from the right-hand side.*

The HE schemes can be classified into fully homomorphic encryption and partially homomorphic encryption [246], as follows.

• *Partial Homomorphic Encryption (PHE):* PHE schemes are the earliest type of HE scheme and only support homomorphic addition or homomorphic multiplication [244]. Yucel et al. [247] use PHE to solve the location privacy issues in the charging station scheduling, where frequent scheduling could expose the drivers' location data. The authors first hide the drivers' location information and then process the hidden information with PHE. The proposed scheme allows the drivers to join and leave the charging station dynamically.

• *Somewhat Homomorphic Encryption (SHE):* SHE schemes support a finite number of homomorphic addition and homomorphic multiplication [244]. The work in [248] provides an integer-based SHE scheme based on the duration between each data transmission from the sensor and the data packaging method. By using the proposed scheme, the authors improve the efficiency of the SHE algorithm by reducing the encryption consumption of the sensor data. Yu et al. [249] focus on the scenario of an online ride-hailing service that allows a group of passengers to share a vehicle with a minimum aggregate distance. The authors use the SHE scheme with ciphertext packing in such a scenario to calculate encrypted aggregate distance. The ciphertext packing is utilized to ensure that the actual value of the aggregate distance cannot be leaked.

• *Fully Homomorphic Encryption (FHE):* FHE schemes support an infinite number of homomorphic [244], but it has limitations of low efficiency and high computational consumption [250]. To decrease the communication overhead of the FHE scheme, Perma et al. [251] combine FHE with pseudonyms. The proposed method uses FHE to encrypt location data transited in vehicular networks, which also avoids the linkage between different pseudonyms of the same vehicle. By considering that all LBSs have deadline and mobility constraints, Mohammed et al. [252] propose a cost-efficient vehicular fog cloud computing method where the FHE is utilized for security. Farouk et al. [3] utilize a network simulator to simulate urban mobility, and a cloud simulation to simulate the real world and protect LBS queries with FHE based on the advanced encryption standard. The authors outsource encrypted location data to the cloud server so the drivers can securely obtain accurate LBS query results.

Nevertheless, there is no efficient homomorphic encryption-based secure computation method in road networks [244].

Private Information Retrieval

Private Information Retrieval (PIR) enables the drivers to request data items without revealing which item is retrieved [253]–[255]. However, the computational consumption of the PIR could be high that it is difficult to allocate in practice [256]. To overcome this limitation, Tan et al. [256] focus on road networks where the authors apply the transportation information of road networks as prior knowledge of PIR that significantly decreases the computational cost.

Searchable Encryption (SE)

SE schemes are considered as a server-side privacy-enhancing technique, which enables the drivers to keep plaintext keywords safe in searching LBSs [257], [258]. SE schemes are also be combined with other methods or tools to improve precision and efficiency, e.g., group signatures, Cuckoo filter, Pederson commitment, smart contract, public-key encryption, and proxy re-

encryption [259], [260]. However, most existing SE schemes would provide wrong results for geometric range searches [261]. To solve this problem, Chen et al. [262] develop a novel SE method under a public-key system supporting arbitrary geometric area searches. The authors avoid the false positive result by introducing inner product encryption. The developed method can achieve 100% accuracy of searching results.

2.2.3 User-server-interface LPPMs

Channel-side techniques encompass mechanisms that act on the security and privacy properties of the set-up communication channel between the server and end-users [263]. Secure communication and trust in the third party are popular in the user-server-interface LPPMs.

Secure Communication

Secure communication can be classified as follows.

• *Encrypted Communication Protocol:* To avoid pervasive communication surveillance, it is essential to emphasize that encrypted channels need to be implemented and configured correctly [264]. In the LBSs of vehicular networks, traditional protocols are widely used [265], e.g., the Transport Layer Security protocol (TLS) [266], [267] and the Secure Shell (SSH) protocols [268], [269]. TLS and SSH rely on public key cryptography techniques so drivers and servers can set up an encrypted channel without sharing secrets [270]. Nevertheless, the traditional protocols still have limitations. For example, SSH requires that the end-users can perform periodical verification [271], which is resource-consuming [272].

• *End-to-end Encrypted Services:* End-to-end encrypted services prevent the adversaries from accessing the location data when transferring among end-users, meaning that only the intended receiver can decrypt the data [273]. End-to-end encrypted services are seldom used in vehicular network scenarios currently. However, as future vehicular networks may require the vehicles to cooperate with others to complete a task, end-to-end encrypted services could be widely used in the future.

Trusted Third Party (TTP)

The TTP techniques, which are commonly based on anonymity and pseudonym, are proposed to prevent malicious entities from obtaining the drivers' sensitive information [274]–[276]. A set of malicious entities could be more powerful than the legal drivers in that they could have the capability to trace the drivers and estimate other private information of the drivers [1]. However, the existing studies [277]–[280] prefer to protect location privacy without any trusted third party, which is a strict scenario.

2.2.4 Trade-off between Location Privacy and Data Utility

Location privacy and data utility are two competing objectives of vehicular networks. The data utility is one of the major concerns in location privacy preservation to measure how the privacy-preserving methods influence LBSs. The drivers are increasingly wary of their location privacy, which may discourage their location information sharing with the LBS platforms. For LBS platforms, they desire the actual locations of the drivers for commercial purposes. The

over-protection will diminish the data utility of the LBSs [281]. Therefore, a privacy-preserving method should maximize data utility with the best privacy-preserving capability.

Many LPPMs focus on the trade-off between location privacy and data utility, as shown in Table 2.3. There are six major methods to achieve the trade-off between location privacy and data utility, as follows.

Blockchain

Blockchain can be used to realize distribution management. LPPMs can access the required information if the information is stored in the blockchain [288]. This technology can decrease time delays and detect the malicious vehicles. Luo et al. [17] use blockchain to store the historical trust information of vehicles and employ Dirichlet distribution to allow vehicles to cooperate with others by utilizing anonymity. The authors reduce the communication delay with blockchain-based trust management. Li et al. [229] develop two metrics, i.e., connectivity and average distance, to measure the data utility of k-anonymity. With these two metrics, the authors evaluate the developed framework, which uses blockchain to solve trust management during the exchanging of safety beacon messages. The experimental results show that the blockchain-based k-anonymity framework decreases the data processing time.

Malik et al. [14] avoid the dependency on a trusted third party by developing a blockchain-based authentication and revocation method which decreases the consumption of computation and communication. Wang et al. [201] combine the blind signature with blockchain to achieve a scheme that has high effectiveness and applicability. The authors use blockchain to store the public key of the vehicle, which can be used for authentication by comparing the calculated Merkel's root value. Tang et al. [260] manipulate group signature and multiple tools when designing a blockchain-based privacy-preserving scope-query searchable encryption scheme, which achieves fairness and accurate parking lot sharing. Chaudhary et al. [282] also use blockchain to overcome the limitations of k-anonymity, while Liang et al. [286] create cloaking region according to the blockchain. Wang et al. [287] evaluate the trustworthiness of vehicles by using blockchain-based RSUs, and Lu et al. [283] use the proofs of presence and absence in blockchain to protect location privacy. The authors ensure the high validity of the two methods and avoid the leakage of identity. Boualouache et al. [284] and Samuel [285] use blockchain to protect location privacy by preventing semantic, linking, and data mining attacks.

Adaptive Parameter

Some LPPMs achieve the trade-off between location privacy and data utility by tuning the parameters according to the drivers' requirements. The adaptive parameters-based LPPMs can decrease unnecessary resource consumption and data utility loss. Lu et al. [181] define the quantitative measurements of location privacy and perfect privacy. By using these measurements, vehicles can route location data with the assistance of other vehicles to obtain efficient, secure and accurate LBSs. Shahandashti et al. [185] and Kaaniche et al. [191] permit the vehicles to only reveal required data based on their attributes and requirements, which can achieve high efficiency and data utility without breaching the anonymity of the vehicles. Ma et al. [15], Zhou et al. [212], Li et al. [213], and Li et al. [214] develop the location privacy-preserving mechanisms which can adaptively set the parameters based on the drivers' protection and utility requirements. Li et al. [218] allow the vehicles to swap their pseudonyms according to their driving status, which can achieve high privacy protection and data utility. Zhu et al. [238] ensure the

Table 2.3: Existing LPPMs in the trade-off between location privacy and data utility.

Methods	**Techniques**	**LPPM category**	**Papers**
Blockchain	Distribution Management	k-anonymity	[17], [229], [282]
		others	[14], [201], [260], [283]–[287]
Adaptive parameters	Tuning parameters	Pass-and-run	[181]
		Attribute-based signature	[185], [191]
		Geo-I	[15], [212]–[214]
		Pseudonym	[218]
		Other anonymity	[238]
		General differential privacy	[240], [241]
		SHE	[248]
		FHE	[252]
Hybrid	Combining multiple methodologies	Group signature	[193], [196]
		Sanitizable signature	[200]
		Dummy-based	[205]
		LDP	[210]
		Joint differential privacy	[243]
		SE	[262]
Encryption optimization	Reducing computational and communication consumption	Group signature	[195]
		Blind signature	[202]
		Secure computation	[5]
		Mix-zone	[236]
		PHE	[247]
		FHE	[251]
Parameter simplification and optimization	Simplifying process	Group signature	[194], [198]
		Blind signature	[202]
		Dummy-based	[206]
		Pseudonym	[217]
		Other anonymity	[239]
		FHE	[3]
		Private information retrieval	[256]

Virtual nodes	Introducing cooperators	Group signature	[197]
		Pseudonym	[217]
		k-anonymity	[235]
		TTP	[277]–[280]

effectiveness of location data utility by using hashmap-based short randomizable signature. By controlling the noise of differential privacy, Ghane et al. [240] and Miao et al. [241] improve the retrieval accuracy and decrease the time complexity to balance the location privacy protection and data utility. Subramaniyaswamy et al. [248] and Mohammed [252] use encrypt location data in the edge of vehicular networks without losing data. The authors reduce the time delay and recourse consumption of encryption.

Hybrid Approach

Some LPPMs utilize different methodologies to improve efficiency. The hybrid LPPMs can obtain great benefits and overcome limitations by mixing various methodologies. Zheng et al. [193] and Mundhe et al. [197] combine group signatures and pseudonyms to achieve identity verification. By using the pseudonym, the signature could be short, and the computational consumption could be reduced. Similarly, work in [200] jointly uses log anonymization and sanitizable signature to mitigate privacy threats. Niu et al. [205] realize k-anonymity by utilizing dummy locations, which considers the side information of the vehicles, and Erlingsson et al. [210] use differential privacy to eliminate the signal at the user side. The two methods can prevent the linking of pseudonyms by the adversaries. With joint differential privacy, Tong et al. [243] research on the private spatial index, private distributed optimization, and private dual decomposition techniques to achieve the balance. Chen et al. [262] use SE and computational private information retrieval to decrease computational consumption, ensuring effectiveness, which considers the restriction of road networks as the prior knowledge.

Encryption Optimization

The encryption methodologies result in high communication consumption, such as bilinear pair cryptography. Some LPPMs focus on optimizing encryption methods to achieve high data utility and low time delay. Wu et al. [195] avoid the high computational cost of message signing and verification by using the elliptic curve cryptosystem rather than bilinear pairs. Sun et al. [202] introduce encryption-based methods, including zero-knowledge verification, one-way hashing, and homomorphic encryption into the partially blind signature to obtain superior effectiveness of network response and privacy preservation. Zhou et al. [5] employ the noncolluding servers, i.e., the cloud and the cryptographic service provider, to skip the interaction between the vehicle and the LBS servers. The authors eliminate redundant encrypted LBS data before authentication, i.e., ciphertext re-encryption, that achieves practicability and protection. Li et al. [236] suggest the vehicles to transmit encrypted data rather than being silent in the mix zone. The authors point out that encryption can decrease the storage consumption of pseudonym management, and increase privacy protection without breaching data utility. Yucel et al. [247] protect location privacy by utilizing homomorphic encryption into bichromatic mutual nearest neighbor assignments, which protect location privacy with low recourse consumption and con-

vergence times. Prema et al. [251] point out that FHE can lead to communication overhead problems. The authors utilize pseudonyms with homomorphic encryption to control the message frequency, which keeps the low overhead of the pseudonym with the high security of the homomorphic encryption.

Parameter Simplification and Optimization

The time delay can be significantly decreased when the LPPMs can simplify the process. For example, the non-certificate LPPMs can eliminate the certificate authentication time delay and certificate storage consumption. Hakeem et al. [194] achieve authentication over multiple zones of large-scale BSs by using a single message and short signature with bilinear pairing cryptography and short-size signature. The authors significantly decrease the generation and verification time of the signature. Mei et al. [198] use a certificateless aggregate signature scheme with full aggregation technology to reduce resource consumption. The trade-off between privacy protection and data utility is achieved by considering random oracles under the computational Diffie-Hellman assumption. Sun et al. [202] optimize a fog-bus-based vehicle crowdsensing framework that severs the relationship between the identity and location data. The authors simplify the data process in data reporting, reputation management, and reward issuing so that the effectiveness can be improved.

Liu et al. [206] focus on the time reachability, direction similarity, and in-degree/out-degree of the location data. The developed spatiotemporal correlation-aware LPPM simplifies neighboring location sets to achieve personalized location privacy protection. Wang et al. [217] improve the data utility of pseudonyms by introducing a trigger-based structure, avoiding the frequent pseudonym changing. Singh et al. [239] simplify the required characteristics of the vehicles by using location and speed instead of their identities. The authors allow each vehicle to send data of others masquerading as its own location data, which significantly decreases the traceability of the pseudonym-base LPPM without breaching the data utility. Farouk et al. [3] outsource the location data to a cloud server that prevents the vehicles from sharing their location data with multiple different entities. By only sharing data with the cloud server, the vehicle can obtain LBSs with a low delay. Tan et al. [256] use the prior knowledge of road networks to decrease the consumption of computational private information retrieval in preprocessing and communication.

Virtual Node

Some LPPMs, e.g., k-anonymity, need a specific number of neighbors to provide a high privacy-preserving capability. The LPPMs can generate virtual nodes to act as physical vehicles to overcome this limitation. The virtual nodes occupy limited storage space. Thus, the privacy-preserving capability can be improved without decreasing the data utility. Mundhe et al. [197] avoid the trusted third party in authentication by employing a lattice-based ring signature. The scheme provides a low authentication delay and cost, which improves the effectiveness of the data utility to the signature-based LPPMs. Wang et al. [217] introduce virtual devices, i.e., triggers, to assist pseudonym changing, which improves the privacy protection of pseudonym-based LPPMs. Zhu et al. [238] distribute fixed mixing zone in road networks to avoid the adversaries linking the pseudonyms. The presented method provides stable location privacy preservation, which is not influenced by the number of vehicles in the mixing zone.

Table 2.3: Trade-off between location privacy and data utility of in-vehicle tracking in practice: COVID-19 applications.

Architecture	**Applications**	**Required information**	**Limitations**
Centralized	COVIDSafe (Australia) [289]	Encounter information	The personal information of the driver, e.g., identity, phone number, and email. Encounter information that is frequently recorded exposes the driver's trajectories. Private information can be inferred from the reported geographical information. The geographical information is disclosed to the public for epidemic prevention
	E-Tabib (Azerbaijan) [290]		
	BeAware Bahrain (Bahrain) [291]		
	Corona Tracer BD (Bangladesh) [292]		
	Taiwan Social Distancing (China) [293]		
	TousAntiCovid (France) [294]		
	VirusRadar (Hungary) [295]		
	Rakning C-19 (Iceland) [296]		
	Smittestopp (Norway) [297]		
	BlueTrace (Singapore) [298]		
	Alipay (China) [299]	Geographical information	
	WeChat (China) [300]		
	LeaveHomeSafe (China) [301]		

Decentralized	Stopp Corona (Austria) [302]	Encounter information	The malicious drivers can infer the trajectories of patients and match the trajectories with identifiers by creating multiple accounts and recording multiple routes. The geographical information is disclosed to the public for epidemic prevention
	Stop COVID-19 (Croatia) [303]		
	eRouška (Czechia) [304]		
	Smittestop (Denmark) [305]		
	Koronavilkku (Finland) [306]		
	Corona-Warn-App (Germany) [307]		
	Apturi Covid (Latvia) [308]		
	Radar COVID (Spain) [309]		
	SwissCovid (Switzerland) [310]		
	NHS COVID-19 (United Kingdom) [311]		
	DP-3T (European) [312]		
	PACT (USA) [313]		
	Private Kit: Safe Path (USA) [314]	Geographical information	
	South Korea system [315]		
	HaMagen (Israel) [316]		

In summary, existing methods that balance location privacy and data utility cannot perfectly satisfy the future 5G/6G-enabled vehicular networks. For example, the Geo-I-based LPPMs inject controlled noise in location data when the data is reported to the LBSs. Nevertheless, inaccurate location data will introduce extra data optimization processes, which degrade the consumption of big data analysis in 5G/6G networks [317]. The data utility will be significantly decreased even if the Geo-I-based methods finely tune the adaptive parameters. This is unacceptable to 5G/6G service providers. The virtual nodes-based method needs to provide the expected balance between location privacy-preserving capability and data utility. The reason is that the future 5G/6G-enabled vehicular networks will deploy numerous sensors to monitor the environment so that virtual nodes can be eliminated. Thus, the challenge of balancing location privacy and data utility will become serious as the requirements of high-precision and novel technologies are proposed in future 5G/6G vehicular networks.

Example of Localization Application

Different from the theoretical trade-off between location privacy and data utility, the applications in practice are more interested in data utility. According to Elbir et al. [318], it is highlighted that vehicular networks offer potential applications in tracking the spread of diseases, identifying instances of social distancing non-compliance, and facilitating health-related services. This is attributed to the widespread use of public and private transportation by individuals in their daily lives. With the sensors equipped on vehicles, the encounter history with social distance can be collected and sent to the authorities [318]. The usage of vehicular communication can improve the detection of the infection, which can prevent the spread of the disease. However, in the pandemic scenario, e.g., COVID-19, tracking also brings challenges. Location data collection and exchange in the fight against COVID-19 is an excellent example of in-vehicle tracking. Existing applications for epidemic prevention can be classified into centralized and decentralized management. In centralized management cases, the medical information of confirmed cases is monitored by governments or institutions. Governments or institutions will only disclose the trajectory of the patient when a new patient appears. In decentralized management cases, the location data is managed by the drivers. Only if the drivers are infected, their trajectories are exposed to authorities. Strict management of such information can save the world from the virus, but the location privacy of the patient is exposed. Table 2.3 shows the benefits and limitations of related applications. In Table 2.3, the existing COVID-19 applications are classified into centralized and decentralized applications.

The centralized applications (e.g., TraceTogether (Singapore) [319], COVIDSafe (Australia) [289], and BlueTrace (Singapore) [298]) record the encrypted encounter history rather than the location information. Only the authority can decrypt the encounter history. The driver's identification can be obtained by the authority if the driver is infected. Other drivers can check their risk of infection by the encounter history. However, the patients' identities and trajectories are exposed to the authority, which increases the risk of location privacy disclosure. The privacy-preserving capability of the applications can be decreased if other drivers collude to infer targets.

DP-3T (European) [312] and PACT (USA) [313] are two examples of decentralized applications. The two applications create a periodical key to generate several ephemeral identifiers. The ephemeral identifiers are broadcasted as a beacon message within a region. The two applications store the received beacon messages with extra information. The applications match the ephemeral identifiers based on the information from the authority. Nevertheless, the mali-

cious drivers can infer the trajectories of patients and match the trajectories with identifiers by creating multiple accounts and recording multiple routes. Private Kit: Safe Path (USA) [314] and South Korea system [315] are the other types of COVID-19 applications. They periodically record the drivers' trajectories with time-stamped log and pseudonyms based on the GPS (the South Korean system record detailed personal information). The drivers can report their trajectories to the authority if they are infected. The authority can select information to be exposed to the public. The trajectories are disclosed to the public for epidemic prevention. Nevertheless, the applications cannot offer the expected location privacy-preserving capability, as everyone can view the trajectories.

The existing COVID-19 contract tracing applications balance location privacy protection capability and data utility by using anonymity [320]. The applications, e.g., TraceTogether (Singapore) [319] and COVIDSafe (Australia) [289], periodically broadcast random time-varying tokens as the driver's temporary IDs [321]. The applications record the encounter information of the drivers, which are employed to report infection risk [322]. However, the authorized entities can still link the temporary IDs and the driver's personal information (e.g., trajectory, phone number), even if the applications do not collect personal information on purpose [321].

2.3 Conclusion

In this chapter, we highlighted the threats of the existing localization techniques to location privacy protection. We then reviewed the existing LPPMs and discussed their advantages and limitations. The methods to balance the data utility and location privacy protection were researched in theoretical and practical scenarios.

The existing works in the field of location privacy preservation in vehicular networks have overlooked several crucial aspects. Firstly, many traditional methods primarily focus on generic 2D planar spaces and neglect the unique intricacies of road networks. These oversights result in impractical outcomes, often obfuscating locations that are off-road and thus irrelevant to vehicular navigation and data sharing. Furthermore, these approaches tend to fall short in providing effective privacy protection for drivers in real-world scenarios, where location privacy requirements can vary significantly depending on the context. Existing methods often lack the adaptability required to meet these varying privacy needs. Another critical gap in the existing research is the limited consideration of differential privacy, specifically in the context of road networks. The additive nature of differential privacy has been overlooked, making it challenging to detect illegal trajectories accurately. Detecting and mitigating malicious or unauthorized vehicle trajectories is essential in ensuring the reliability and security of vehicular networks. Furthermore, prior research has not sufficiently explored cooperative privacy protection mechanisms that leverage the collective power of multiple vehicles. Cooperative approaches have the potential to enhance privacy while maintaining data utility effectively. These collaborative methods can provide a higher level of protection against privacy breaches and collusion attacks. The integration of pseudonym swapping with obfuscation techniques remains underutilized in the existing literature, despite its potential to improve privacy preservation. In summary, this thesis aims to address these research gaps by proposing a comprehensive differential privacy framework tailored to road networks, personalized location privacy-preserving mechanisms, and cooperative privacy protection strategies, ultimately advancing the state-of-the-art in location privacy for vehicular networks.

Chapter 3

Personalized Location Privacy with Road Network-Indistinguishability

3.1 Introduction

Obfuscation-based schemes can protect location privacy against untrusted LBS servers and eavesdroppers [72]–[74]. The idea is based on the local differential privacy technology where users locally perturb raw data with differential privacy schemes and employ the obfuscated data rather than the accurate data [75]. In obfuscation-based schemes, drivers generate obfuscated locations by adding controllable noise in their actual locations and report the obfuscated locations in LBS requests [69]. The obfuscated locations are selected according to the distances between the obfuscated and actual locations under the constraint of indistinguishability [74]. Although the obfuscation-based schemes compromise location accuracy, they can still be used in location-insensitive LBS, such as location-based recommendations [76].

Obfuscation-based schemes have not been rigorously developed for vehicles in road networks. In general-purpose obfuscation schemes, e.g., two-dimensional (2D) Laplace location privacy-preserving schemes, the distance between an actual location and the corresponding obfuscated location is measured using the Euclidean distance [73], [74], [77]. This, however, underestimates the distance between two locations on road networks because the route distance is no shorter than the Euclidean distance between any two locations [78]. Another issue of general-purpose obfuscation schemes is that they may generate off-road obfuscated locations, for example, on a river [79]. Additional steps can be introduced to avoid off-road locations, which, however, increase the computational complexity [80].

The privacy-preserving levels of locations are expected to be fine-tuned across road networks to balance data utility and location-privacy protection [81]. Drivers can be sensitive to a small number of locations in road networks. If all locations are obfuscated at the highest level as the sensitive locations, the insensitive locations are over-obfuscated, and the data utility is penalized [82]. If only the sensitive locations are obfuscated, the sensitive locations are statistically different from the insensitive locations. This compromises location privacy [83].

In this chapter, we propose a Personalized Location Privacy-Preserving (PLPP) scheme which provides customized location-privacy protection for vehicles in road networks. We start by defining Road Network-Indistinguishability (RN-I), where the distinguishability of locations is measured by the route distance. A connection-interval obfuscation algorithm is proposed

to select on-road obfuscated locations according to privacy budgets and the route distances between obfuscated and actual locations. A personalization algorithm is designed to customize the privacy budgets of connections based on the sensitive locations specified by drivers.

Table 3.1: Summary of notations and abbreviations

Notation	Description
RN-I	Road Network-Indistinguishability
PLPP	Personalized Location Privacy-Preserving
Geo-I	Geo-Indistinguishability
$\mathcal{G}$	The graph transformed from a road network
$\mathcal{V}$	The set of connections on $\mathcal{G}$
$\mathcal{E}$	The edge set on $\mathcal{G}$
p/p'	The actual/obfuscated location
v/v'	The actual/obfuscated connection
$d_{\mathcal{G}}(p, p')$	The shortest route distance between p and p'
ϵ	Privacy budget
u_i	The i-th interval
α	The number of the intervals in a road
D	Connection degree
G_i	The i-th segment of $\mathcal{G}$
v_i^m	The i-th m-hop connection from sensitive locations
$W_{v_i^m}/W_{G_i}$	The weight of v_i^m/G_i
SL	The set of sensitive locations
sl_i	The i-th sensitive location
AEs	Adversary estimation Errors
r	The radius of the obfuscation region
δ	Weight threshold

The key contributions of our work are listed as follows.

1) We define a new measure of location indistinguishability for road networks by applying the concept of differential privacy. The indistinguishability is measured only on road networks (as opposed to continuous 2D space). Route distances (i.e., distances along roads) are used to derive the differential privacy upper bound. Off-road locations are precluded from obfuscation operations.

2) We propose a new dual-obfuscation algorithm that firstly probabilistically obfuscates an actual location to a connection and then obfuscates the location into a road interval between the actual location and the obfuscated connection. By carefully designing the probabilities, we prove that the dual-obfuscation design satisfies the new differential privacy-based definition of road network-indistinguishability.

3) Considering non-uniform location privacy requirements of a driver at different locations, we apply the nearest neighbor interpolation to specify the privacy budgets for all locations based on the sensitive locations of the driver. The locations of the vehicle can be obfuscated consistently without exposing the sensitive locations, while improving the utility of the location data of less sensitive locations.

We conduct comprehensive experiments on two real-world trajectory datasets and compare the

proposed PLPP scheme with the existing 2D Laplace location privacy-preserving schemes [73], [77], [89]. Experimental results show that the PLPP scheme outperforms the 2D Laplace schemes in terms of data utility, privacy-preserving level, and efficiency.

The rest of this chapter is organized as follows. In Chapter 3.2, related works are reviewed, followed by the system model in Chapter 3.3. The proposed privacy-preserving scheme is presented in Chapter 3.4 and experimentally evaluated in Chapter 3.5, followed by conclusions in Chapter 3.6. Notations and abbreviations used in this chapter are listed in Table 3.1.

3.2 Related Work

Studies about cryptography, signature, trust mechanisms, LBS design, and game-theoretical models have been carried out to protect location data [323]. The scheme developed in [324] protects both past and current LBS queries from each user with modified linkable spontaneous anonymous signature and the oblivious transfer technique. The scheme presented in [325] is an asynchronous localization protocol that constructs the relationship between propagation delay and location to eliminate the side-effect of asynchronous clock and mobility. Location-based recommendation is a popular LBS application and requires three types of information, i.e., user, activity, and location [326]. The location-aware recommendation scheme developed in [327] uses locality-sensitive hashing to hide the location information of both users and services. The scheme can provide accurate recommendation results while preserving users' location privacy.

Various schemes have been proposed to protect location privacy for vehicles, such as mix-zone [328], k-anonymity [329], and obfuscation [330]. In mix-zone schemes, drivers stop sending LBS queries in *mixing* areas and change their pseudonyms when leaving the areas [328]. Nevertheless, the adversary can link pseudonyms by analyzing the timing information and transition probabilities of the drivers. In k-anonymity schemes, drivers send a group of locations, including actual locations and phony locations, to LBS servers [329]. The k-anonymity schemes have a high communication cost because the drivers need to communicate with the k nearest neighbors. Obfuscation-based schemes use perturbed locations rather than actual locations in location-insensitive LBS, such as location-based recommendations [331]. Drivers can select desired results from LBS returns based on their actual locations [332].

Table 3.2: Related Obfuscation-based Schemes

Ref.	Scenario	Methodology	Personalized
[74], [77]	2D	Geo-I	-
[73]	2D and RN	Geo-I	-
[69]	RN	Geo-I	-
[89]	2D	Geo-I	Yes
PLPP	RN	RN-I	Yes

The obfuscation-based schemes are derived from differential privacy (DP) [333], which aims to protect an individual's private information while publishing aggregated data about a dataset or a single message [240], [334]. Popular obfuscation-based works are listed in Table 3.2, where we compare the scenarios and the methodologies of the schemes.

Based on differential privacy, Andrés *et al.* develop a scheme called Geo-Indistinguishability (Geo-I) [74]. The authors employ the Laplace scheme [335] to realize Geo-I on a two-dimension

(2D) plane. The Geo-I properly reports obfuscated locations surrounding the actual location based on the obfuscation probability distribution when a driver requests LBS. The results show that the Geo-I can achieve effective location privacy protection under infrequent location updates [336]. The scheme developed in [77] reduces the computational consumption of the Geo-I on the 2D plane by introducing additional servers. A driver in the region covered by the additional servers only calculates obfuscated locations once and repeatedly reports the obfuscated location, so that the average computational consumption is almost zero. The scheme in [69] reduces the computational consumption of the Geo-I by dividing roads into intervals with a certain length. The locations in the same interval can be obfuscated in the same way, hence reducing the computational consumption.

Some obfuscation-based schemes focus on location privacy in road networks. The Graph-Exponential scheme (GEM), which satisfies the Geo-I, evaluates the privacy protection level and data utility of the traditional Geo-I in road networks [73]. The GEM sets the connections (such as turns, forks, and intersections) in a road network as the obfuscation candidates. The scheme obfuscates drivers' actual locations to connections directly based on the Geo-I. The scheme in [69] discretizes a road network with the same length intervals. The authors employ the route distance between two intervals to measure the indistinguishability of the Geo-I. Nevertheless, it is challenging to set the same length intervals in the road network. If the intervals are short to accommodate short roads, the computational consumption increases [334]. If the intervals contain several short roads, the correlation between the connections and privacy may be ignored.

Various schemes protect drivers' sensitive locations at different levels to realize personalization. The personalized scheme in [337] measures the privacy requirements with personal attributes, e.g., access duration, frequency, and regularity. The scheme formulates an incomplete information game to balance quality-of-service and privacy protection. The movement regularity-based privacy requirement is studied in [338] for personalized pseudonym swapping. The scheme in [339] measures privacy requirements using intimacy which specifies community edge density in social networks. Differential privacy and generative adversarial networks are employed to add noise to raw data. The scheme [281] specifies the personal privacy requirements of a location to be negatively correlated with the number of hops to sensitive locations. The algorithm in [340] orchestrates semantic privacy and location privacy based on drivers' requirements, which are measured according to the relationship between drivers. A game-theoretic model is then constructed in coupling with social-distance-based differential privacy to protect location privacy. The scheme in [89] designs the privacy requirement to be negatively correlated with the Euclidean distance between the current location and the last inferred location. The privacy requirement is used to calculate the privacy budget for obfuscation. This design reduces the exposure probability but requires the real-time calculation of the privacy requirement.

To the best of our knowledge, none of the existing personalized obfuscation schemes has considered the features of road networks, such as route distance and road network topology. Moreover, most of the existing personalized schemes require heavy computation based on historical behaviors. The proposed methodology aims to safeguard the location privacy of drivers within road networks. The PLPP scheme comprises two key components: a Connection-Interval Obfuscation algorithm and a Personalization algorithm. The Connection-Interval Obfuscation algorithm discretizes road edges into intervals and obfuscates intervals, rather than individual locations, to reduce computational overhead. This dual-obfuscation process, consisting of connection and interval perturbations, ensures location privacy while enhancing data utility. The Personalization algorithm empowers drivers to specify sensitive locations, allowing for varying levels of location privacy protection based on personalized privacy requirements. By calculating

connection weights in a hop-by-hop manner, the algorithm customizes privacy budgets across the road network, ensuring sensitive locations receive the highest protection, while optimizing data utility for less sensitive areas. The PLPP scheme adheres to the proposed (ϵ,r)-Road Network-Indistinguishability, effectively safeguarding location privacy near sensitive locations and providing robust privacy protection in road network-based location-based services.

3.3 System Model

In this section, we illustrate the road network and adversary model, and the important concept of this chapter.

3.3.1 Road Network Model

A road network can be viewed as a weighted directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of connections (such as turns, forks, and intersections) and $\mathcal{E}$ is the edge set. Nodes in graph $\mathcal{G}$ are located at the same geographic positions as in the real world. The weight of an edge is the length of the corresponding road.

A vehicle can only move along an edge e, following the traffic rules at any given time. Drivers can access the LBS by sending their location information to the LBS servers. We consider the case that the LBS servers are not trusted and may disclose the location records of drivers to adversaries. A driver can conceal its actual location, i.e., p_0, by reporting an obfuscated location, i.e., p'_0, using location obfuscation schemes.

3.3.2 Shift Distance

In this chapter, the definition of the data utility follows that in differential privacy, i.e., the difference between the obfuscated and actual location data, and measures the usability of the obfuscated location [10], [14], [36]. The data utility is higher, i.e., a better LBS response is provided, if the obfuscated location data used in the corresponding LBS request is closer to the actual location. In this sense, shift distance can be used to quantify the data utility of location obfuscation schemes. The smaller a shift distance is, the more accurate the LBS is and the higher the data utility is. Most existing location obfuscation schemes, such as [77] and [89], are designed for 2D planes (or maps) and employ the Euclidean distance between an actual location and its obfuscated location as the shift distance. In this chapter, we consider location obfuscation schemes in road networks and define the shift distance as

Definition 4 (Shift Distance)**.** *Given an actual location p_0 and its corresponding obfuscated location p'_0, the shift distance is the shortest route distance between the actual location p_0 and the obfuscated location p'_0.*

The data utility decreases with the increasing distance between an actual location and its obfuscated version, i.e., the shift distance. The shift distance has been extensively used in existing literature (e.g., [69] and [77]) as the only metric for evaluating the data utility.

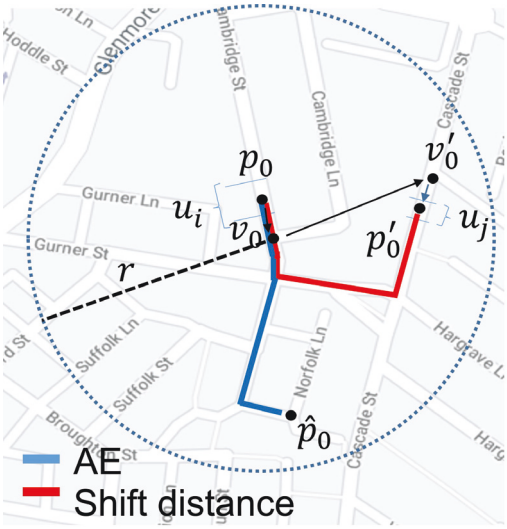


Figure 3.1: An example of the connection-interval obfuscation of the PLPP scheme.

3.3.3 Adversary Model

Adversaries in this chapter are passive internal attackers or external attackers. The adversaries aim to infer the actual locations of drivers from the obfuscated locations in LBS requests. The adversaries are assumed to have full knowledge of drivers' obfuscation scheme and drivers' prior trajectories [69]. With an obfuscated location p_0' of a driver, the adversaries can derive the probability that the driver is actually at p_0 using the Bayes' Theorem, as given by

$$\Pr[p_0|p_0'] = \frac{\Pr[p_0'|p_0]\Pr[p_0]}{\int_{p \text{ in } \mathcal{G}} \Pr[p_0'|p]\Pr[p]dp}, \tag{3.1}$$

where p is a location in graph $\mathcal{G}$, and $\Pr[p]$ is the probability that the driver is at location p. $\Pr[p_0'|p]$ is the probability that the driver is actually at p and generates obfuscated location p_0' based on the obfuscation scheme.

For any location p_1 in $\mathcal{G}$, the adversaries can estimate how likely p_1 is the actual location, as given by

$$\Pr[p_1|p_0'] = \frac{\Pr[p_0'|p_1]\Pr[p_1]}{\int_{p \text{ in } \mathcal{G}} \Pr[p_0'|p]\Pr[p]dp}. \tag{3.2}$$

Then, the adversary selects the location $\hat{p}_0$, which is the most likely to be the actual location p_0, as the driver's actual location, i.e.,

$$\hat{p}_0 = \arg\max_{p_1 \text{ in } \mathcal{G}} \Pr[p_1|p_0']. \tag{3.3}$$

Adversary estimation Error (AE) defines the shortest route distance between p_0 and $\hat{p}_0$, and evaluates the privacy protection effect of obfuscation-based schemes. As shown in Fig. 3.1, the route distance between p_0 and p_0' is the shift distance, and the route distance between p_0 and $\hat{p}_0$ is the AE.

3.3.4 Road Network-Indistinguishability

The Geo-I scheme, originally proposed for 2D Euclidean space, can be developed to protect location privacy in weighted and directed road networks. The Geo-I scheme perturbs actual

locations using the concept of differential privacy, where a set of actual locations are geo-indistinguishable to the adversaries. We define the Road Network-Indistinguishability (RN-I) to measure on-road privacy-preserving schemes.

Definition 5 (Road Network-Indistinguishability)**.** *A scheme satisfies (ϵ, r)-Road Network-Indistinguishability if and only if, for an obfuscated location p', any location pair p_i and p_j in the road network ($d_{min}(p_i, p_j) \leq r$) have*

$$\frac{\Pr[p_i|p']}{\Pr[p_j|p']} \leq e^{\epsilon d_{\mathcal{G}}(p_i,p_j)} \frac{\Pr[p_i]}{\Pr[p_j]}, \tag{3.4}$$

where $d_{\mathcal{G}}(p_i, p_j)$ and $d_{min}(p_i, p_j)$ are the shortest route distance and the Euclidean distance between p_i and p_j, respectively. r is the radius of the obfuscation region $\mathbb{R}^2$, the center of which is p_i. ϵ is the privacy budget as used in differential privacy [77]. A low value of ϵ corresponds to a high privacy-preserving capability.

By substituting (3.2) into (3.4), the latter can be transformed as

$$\Pr[p'|p_i] \leq e^{\epsilon d_{\mathcal{G}}(p_i,p_j)} \Pr[p'|p_j], \tag{3.5}$$

which indicates that p_i and p_j are RN-Indistinguishable, for any pair of p_i and p_j.

The route distance measures the shortest distances between two mutually accessible locations. The use of route distance does not introduce extra complexity to the proposed scheme. One reason is that the route distance can be pre-calculated and stored in vehicles. Another reason is that the route distances can help preclude off-road locations, hence reducing the computational complexity required for obfuscation.

3.4 Personalized Location Privacy-Preserving Scheme

Applying the RN-I, we propose the Personalized Location Privacy-Preserving (PLPP) scheme to protect the location privacy of drivers in road networks. The PLPP scheme consists of a location obfuscation algorithm and a personalization algorithm. The obfuscation algorithm perturbs actual locations using the privacy budget calculated by the personalization algorithm.

3.4.1 Connection-Interval Obfuscation

There are two major types of road division schemes: (a) roads are divided into the same number of intervals, as shown in Fig. 3.2(a); and (b) the intervals in the road network have the same length, as shown in Fig. 3.2(b).

The schemes (a) and (b) can be used to discretize and segment road networks. We choose method (a), because it can readily and meaningfully obfuscate a vehicle's location to somewhere on a different yet connected road. This obfuscated location preserves location privacy, while remaining relevant for typical location-based services. The resolution (or granularity) of the segmentation is on a road basis. It is adequate, and consistent with the requirements of the location privacy and the location accuracy for typical location-based services. While method (b) can also work under the proposed PLPP scheme, the method would excessively

(a) (b)

Figure 3.2: Roads are divided into intervals by (a) the same number of intervals; (b) the identical length of intervals.

segment a typical urban road network. The granularity of the method could be excessively fine, as compared to our goal of obfuscating a vehicle's location to a different yet connected road. Furthermore, it could be tedious and computationally expensive to divide an urban road network into segments with equal length. This is due to the fact that the lengths of different roads can vary dramatically in a road network. The largest common factor of the road distances can be very short, resulting in an excessively large number of segments. It is worth pointing out that the method described in (b) can offer practical values on highways with long distances between entrances/exits. By adopting the method, the proposed PLPP scheme can be applied to highways and obfuscate vehicles into different segments of a long highway.

The PLPP scheme discretizes every edge into intervals and obfuscates intervals rather than single locations to reduce computational consumption. We evenly divide each edge into α intervals (α is a certain number). The interval index can start from either side of the edge because interval u_i is obfuscated in the same way as $u_{\alpha-i}$.

Vehicles perturb actual locations by locally running the connection-interval obfuscation algorithm in the PLPP scheme, including a connection perturbation and an interval perturbation. The connection perturbation determines the privacy-preserving level of the PLPP scheme, while the interval perturbation obfuscates connections to intervals to improve data utility and enlarge AE. The pseudocode of the connection-interval obfuscation algorithm is shown in Algorithm 1. An example is illustrated in Fig. 3.1.

In Algorithm 1, steps 4 to 11 are the connection perturbation. An actual location p_0 in the interval u_i is first mapped to the nearest adjacent connection of u_i, denoted by v_0. Then, v_0 is perturbed to another connection v_0' within the obfuscation region $\mathbb{R}^2$ (centered at v_0 with radius r) based on the connection perturbation probability, i.e., $\Pr[v_0'|v_0]$, as given by

$$\Pr[v_0'|v_0] = \frac{1}{\sum\limits_{v \,\text{in}\, \mathbb{R}^2} e^{-\frac{\epsilon}{2} d_{\mathcal{G}}(v_0, v)}} e^{-\frac{\epsilon}{2} d_{\mathcal{G}}(v_0, v_0')}, \tag{3.6}$$

where v is a connection in the obfuscation region $\mathbb{R}^2$.

The interval perturbation is given by steps 12 to 22 in Algorithm 1. Firstly, the PLPP scheme selects the connection v_1 that is adjacent to v_0' and along the shortest route between p_0 and v_0'. Then, the scheme selects an interval u_j on road (v_0', v_1) according to the interval perturbation probabilities, i.e., $\Pr[u_j|u_i]$, as given by

$$\Pr[u_j|u_i] = \frac{1}{\sum\limits_{k=1}^{\alpha} e^{-\frac{|i-k|}{2\alpha}\epsilon}} e^{-\frac{|i-j|}{2\alpha}\epsilon}, \tag{3.7}$$

Algorithm 1 Connection-Interval Obfuscation Algorithm.

Input:
The road network, $\mathcal{G}$;
The segment privacy budget, ϵ;
An actual location, p_0;
The number of intervals, α;
The obfuscation radius of each connection, r.

Output:
The obfuscated location, p'_0.

▷ **Setup**
$u_i \leftarrow (\mathcal{G},\ p_0)$ ▷ p_0 is in u_i
$v_0 \leftarrow (\mathcal{G},\ u_i)$ ▷ v_0 is u_i's nearest adjacent connection
$\mathbb{R}^2 \leftarrow (\mathcal{G},\ v_0,\ r)$ ▷ identify the obfuscation region $\mathbb{R}^2$
▷ **Connection perturbation**
for v_i in $\mathbb{R}^2$ **do**
    Calculate $e^{-\frac{\epsilon}{2} d_{\mathcal{G}}(v_0, v_i)}$
end for
Calculate $\sum_{v \text{ in } \mathbb{R}^2} e^{-\frac{\epsilon}{2} d_{\mathcal{G}}(v_0, v)}$
for v_i in $\mathbb{R}^2$ **do**
    Calculate $\Pr[v_i|v_0]$ ▷ use (3.6)
end for
$v'_0 \leftarrow (\mathbb{R}^2, \Pr[v_i|v_0])$ ▷ Select v'_0 based on $\Pr[v_i|v_0]$
▷ **Interval perturbation**
$v_1 \leftarrow (\mathcal{G},\ p_0,\ v'_0)$ ▷ v_1 is adjacent to v'_0 and on (p_0, v'_0)
for u_k in road $(v_1, v'_0,)$ **do**
    Calculate $e^{-\frac{|i-k|}{2\alpha}\epsilon}$
end for
Calculate $\sum_{k=1}^{\alpha} e^{-\frac{|i-k|}{2\alpha}\epsilon}$
for u_j in road $(v_1, v'_0,)$ **do**
    Calculate $\Pr[u_j|u_i]$ ▷ use (3.7)
end for
$u'_i \leftarrow ((v_1,\ v'_0),\ \Pr[u_j|u_i])$ ▷ Select u'_i based on $\Pr[u_j|u_i]$
$p'_0 \leftarrow (\mathcal{G},\ u'_i)$ ▷ Randomly select a location in u'_i
return The obfuscated location p'_0.

which indicates that intervals u_i and u_j are (ϵ, α)-Indistinguishable. Finally, the PLPP scheme randomly selects a location p'_0 in the interval u_j as the obfuscated location. In the interval perturbation, only the indexes to the intervals are used so that intervals on different roads with the same index are perturbed following the same interval perturbation probability distribution.

The interval perturbation reduces the shift distance by $d_{\mathcal{G}}(v'_0, p'_0)$ and increases data utility compared to connection-only obfuscation schemes, e.g., the connection perturbation and the scheme in [73]. The shift distance, i.e., $d_{\mathcal{G}}(p_0, p'_0)$, of the PLPP scheme is evaluated by

$$d_{\mathcal{G}}(p_0, p'_0) = d_{\mathcal{G}}(p_0, v'_0) - d_{\mathcal{G}}(v'_0, p'_0). \tag{3.8}$$

The interval perturbation increases the AE and improves the location privacy-preserving effect on top of connection-only obfuscation schemes because adversaries can hardly infer the

perturbed connections with on-road obfuscated locations.

Theorem 1. *The PLPP scheme satisfies the (ϵ,r)-RN-Indistinguishability that any two locations in an obfuscation region are (ϵ,r)-RN-Indistinguishable.*

Proof. The connection perturbation in the PLPP scheme is formulated as (3.6). Let $f(v_0) = \sum_{v \text{ in } \mathbb{R}^2} e^{-\frac{\epsilon}{2} d_{\mathcal{G}}(v_0, v)}$. With the actual location p_0, given another actual location p_1 and its nearest adjacent connection v_1 in $\mathbb{R}^2$, we have

$$\frac{\Pr[v'|v_0]}{\Pr[v'|v_1]} = \frac{f(v_1)}{f(v_0)} e^{\frac{\epsilon}{2}(d_{\mathcal{G}}(v_1, v') - d_{\mathcal{G}}(v_0, v'))}, \tag{3.9}$$

where v_0 is the nearest adjacent connection of the actual location p_0 and v' is the obfuscated connection.

We next prove the triangle inequality of the road network.

If v_1 is on the shortest route between v_0 and v_2, we have

$$d_{\mathcal{G}}(v_0, v_2) - d_{\mathcal{G}}(v_1, v_2) = d_{\mathcal{G}}(v_0, v_1). \tag{3.10}$$

If v_1 is not on the shortest route between v_0 and v_2, by assuming that $d_{\mathcal{G}}(v_0, v_2) - d_{\mathcal{G}}(v_1, v_2) > d_{\mathcal{G}}(v_0, v_1)$, we have

$$d_{\mathcal{G}}(v_0, v_2) > d_{\mathcal{G}}(v_0, v_1) + d_{\mathcal{G}}(v_1, v_2),$$

which means that the route between v_0 and v_2 is longer than the route between v_0 and v_2 through v_1. Then, the shortest rout between v_0 and v_2 should pass v_1. Thus, the assumption is false. If v_1 is not on the shortest route between v_0 and v_2, the following inequality holds

$$d_{\mathcal{G}}(v_0, v_2) - d_{\mathcal{G}}(v_1, v_2) \leq d_{\mathcal{G}}(v_0, v_1). \tag{3.11}$$

Combining (3.10) and (3.11), the triangle inequality in road networks is proven.

In other words, *for any three connections (v_0, v_1, v_2) in a road network, $d_{\mathcal{G}}(v_0, v_2) - d_{\mathcal{G}}(v_1, v_2) \leq d_{\mathcal{G}}(v_0, v_1)$ holds.*

Due to the triangle inequality in road networks $d_{\mathcal{G}}(v_0, v_1)$ is no shorter than $d_{\mathcal{G}}(v_1, v') - d_{\mathcal{G}}(v_0, v')$. Then, we have

$$\frac{\Pr[v'|v_0]}{\Pr[v'|v_1]} \leq \frac{f(v_1)}{f(v_0)} e^{\frac{\epsilon}{2} d_{\mathcal{G}}(v_0, v_1)}. \tag{3.12}$$

By employing the triangle inequality, we have $e^{-\frac{\epsilon}{2} d_{\mathcal{G}}(v_1, v)} \leq e^{-\frac{\epsilon}{2}(d_{\mathcal{G}}(v_0, v) - d_{\mathcal{G}}(v_0, v_1))}$. Therefore,

$$\sum_{v \text{ in } \mathbb{R}^2} \left(e^{-\frac{\epsilon}{2} d_{\mathcal{G}}(v_1, v)} - e^{-\frac{\epsilon}{2}(d_{\mathcal{G}}(v_0, v) - d_{\mathcal{G}}(v_0, v_1))} \right) \leq 0, \tag{3.13}$$

which can be rewritten as

$$\sum_{v \text{ in } \mathbb{R}^2} e^{-\frac{\epsilon}{2} d_{\mathcal{G}}(v_1, v)} - e^{\frac{\epsilon}{2} d_{\mathcal{G}}(v_0, v_1)} \sum_{v \text{ in } \mathbb{R}^2} e^{-\frac{\epsilon}{2} d_{\mathcal{G}}(v_0, v)} \leq 0. \tag{3.14}$$

Based on the definition of $f(v_0)$, we have

$$f(v_1) - e^{\frac{\epsilon}{2} d_{\mathcal{G}}(v_0, v_1)} f(v_0) \leq 0. \tag{3.15}$$

Thus, the following inequality holds,

$$\frac{f(v_1)}{f(v_0)} \leq e^{\frac{\epsilon}{2} d_{\mathcal{G}}(v_0, v_1)}. \tag{3.16}$$

Combining (3.12) and (3.16), we have

$$\frac{\Pr[v'|v_0]}{\Pr[v'|v_1]} \leq e^{\frac{\epsilon}{2} d_{\mathcal{G}}(v_0, v_1)} e^{\frac{\epsilon}{2} d_{\mathcal{G}}(v_0, v_1)} = e^{\epsilon d_{\mathcal{G}}(v_0, v_1)}, \tag{3.17}$$

which satisfies (3.5). As a result, the proposed connection perturbation achieves the RN-I.

The interval perturbation is independent of the connection perturbation. This is because the interval perturbation obfuscates interval indexes rather than the outputs of the connection perturbation. Next, we prove the interval perturbation achieves (ϵ, α)-indistinguishability.

Let u_i denote an actual interval (the i-th of α intervals), and u_j is another interval with the same or different sequence number of u_i. u_k is the perturbed interval. u_i, u_j, and u_k can locate on the same road or on different roads. Let $g(u_i) = \sum_{k=1}^{\alpha} e^{-\frac{\epsilon}{2} \frac{|i-k|}{\alpha}}$. Then,

$$\frac{\Pr[u_k|u_i]}{\Pr[u_k|u_j]} = \frac{g(u_j)}{g(u_i)} e^{\frac{\epsilon}{2}(\frac{|j-k|}{\alpha} - \frac{|i-k|}{\alpha})}. \tag{3.18}$$

As $|a| - |b| \leq |a - b|$ (a and b are two real numbers), we have $|j-k| - |i-k| \leq |i-j|$. Thus,

$$\frac{\Pr[u_k|u_i]}{\Pr[u_k|u_j]} \leq \frac{g(u_j)}{g(u_i)} e^{\frac{\epsilon}{2} \frac{|i-j|}{\alpha}}. \tag{3.19}$$

As $|i-k| - |i-j| \leq |j-k|$ holds, we have $e^{-\frac{\epsilon}{2} \frac{|j-k|}{\alpha}} \leq e^{-\frac{\epsilon}{2}(\frac{|i-k|}{\alpha} - \frac{|i-j|}{\alpha})}$. Therefore,

$$\sum_{k=1}^{\alpha} \left(e^{-\frac{\epsilon}{2} \frac{|j-k|}{\alpha}} - e^{-\frac{\epsilon}{2}(\frac{|i-k|}{\alpha} - \frac{|i-j|}{\alpha})} \right) \leq 0, \tag{3.20}$$

which can be rewritten as

$$\sum_{k=1}^{\alpha} e^{-\frac{\epsilon}{2} \frac{|j-k|}{\alpha}} - e^{\frac{\epsilon}{2} \frac{|i-j|}{\alpha}} \sum_{k=1}^{\alpha} e^{-\frac{\epsilon}{2} \frac{|i-k|}{\alpha}} \leq 0. \tag{3.21}$$

Based on the definition of $g(u_i)$, we have

$$g(u_j) - e^{\frac{\epsilon}{2} \frac{|i-j|}{\alpha}} g(u_i) \leq 0. \tag{3.22}$$

Thus, the following inequality holds,

$$\frac{g(u_j)}{g(u_i)} \leq e^{\frac{\epsilon}{2} \frac{|i-j|}{\alpha}}. \tag{3.23}$$

Combining (3.19) and (3.23), we have

$$\frac{\Pr[u_k|u_i]}{\Pr[u_k|u_j]} \leq e^{\frac{\epsilon}{2} \frac{|i-j|}{\alpha}} e^{\frac{\epsilon}{2} \frac{|i-j|}{\alpha}} = e^{\epsilon \frac{|i-j|}{\alpha}}. \tag{3.24}$$

As a result, u_i and u_j are (ϵ, α)-Indistinguishable. With the obfuscated interval u_k, the adversaries cannot distinguish the intervals or infer the obfuscated connection.

Utilizing the PLPP scheme, the actual location is mapped to its nearest adjacent connection in the region $\mathbb{R}^2$. Hence, when the connection perturbation satisfies the RN-I, locations in the region $\mathbb{R}^2$ are (ϵ, r)-Road Network-Indistinguishable.

□

When an adversary obtains an obfuscated location in an LBS request, the adversary first estimates the perturbed connection. Then, the adversary infers the mapped connection with the obfuscated connection. Finally, the adversary tries to estimate the location with the inferred mapped connection and the obfuscated interval. The PLPP scheme can protect location privacy under this adversarial model since the scheme satisfies the RN-I, as given by Theorem 1.

3.4.2 Personalization Algorithm

Personalized location privacy protection can be achieved by protecting a driver's location data at different levels based on the driver's personal privacy requirements. For a driver, a relatively small number of locations are critical and sensitive, and need to be stringently protected. The other less sensitive locations also need to be protected in order to reduce the statistical difference between sensitive locations and less sensitive locations. On the other hand, if all locations are perturbed to the same extent as the sensitive locations, the utility of perturbed location data is compromised excessively. Thus, the privacy-preserving levels across locations should be carefully designed to protect sensitive locations while improving the location data utility of less sensitive locations.

The proposed PLPP scheme allows the drivers to specify sensitive locations and protects sensitive locations with the lowest privacy budget (i.e., at the highest level). Higher privacy budgets can be configured at less sensitive locations, so that the location data at these locations can provide better utility. To achieve this, we propose a personalization algorithm to tune the privacy budget ϵ of connections. The algorithm controls the perturbation level of the dual-obfuscation process. To be specific, a driver first inputs its sensitive locations to the proposed personalization algorithm for configuring the privacy budgets across the road network. The most sensitive locations have the highest weights and the lowest privacy budgets for the best possible privacy protection. The personalization algorithm calculates the weights of all connections hop-by-hop, starting from the sensitive locations. The algorithm exploits the idea of the nearest neighbor interpolation and captures unique road network features, e.g., route distance and road network topology. Finally, the proposed algorithm divides the road network into segments and calculates the privacy budget for each segment based on the weights of connections in the segment. By using this personalized privacy budget configuration, the proposed dual-obfuscation described in Section 3.4.1 can protect location privacy near sensitive locations as expected by the drivers and prevent exposing the sensitive locations.

The road network is treated as a weighted directed graph $\mathcal{G}$. The road network is divided into g segments, i.e., $\mathcal{G} = \{G_1, G_2, \cdots, G_g\}$. The number and the size of segments can be adjusted based on drivers' privacy requirements. The proposed personalization algorithm first calculates the weights of all connections according to the most sensitive locations specified by a driver. The weight of a segment takes the maximum connection weight in the segment, and the privacy

budget of the segment is the ratio of the initial privacy budget to the weight of the segment. The PLPP scheme uses the privacy budget of a segment as the private budgets of the locations in the segment when obfuscating the locations.

The weights of the connections are calculated hop-by-hop, starting from the sensitive locations. Let $\mathcal{V}_m = \{v_1^m, v_2^m, \cdots\}$ denote the set of m-hop connections, where drivers can reach the i-th m-hop connection v_i^m via at least $(m-1)$ connections starting from a sensitive location. An example of $\mathcal{V}_1$ and $\mathcal{V}_2$ is given in Figs. 3.3(a) and 3.3(b). The weight of v_i^m is denoted by $W_{v_i^m}$.

As shown in Algorithm 2, the calculation of connection weights consists of the following steps.

- **Initialization (Steps 1-4):** The weights of sensitive locations and connections adjacent to the sensitive locations are set to 1. The weights of other connections are set to 0. The weight of 1 implies that the privacy budgets of sensitive locations are set to be initial privacy budget ϵ_0.
- **Weight Calculation (Steps 5-16):** The connection weights are calculated hop-by-hop, e.g., calculating the weights of the connections in $\mathcal{V}_3$ after finishing the calculation on $\mathcal{V}_2$. The weight calculation starts from $\mathcal{V}_2$ and ends when weights of all connections have been calculated or all the connection weights in $\mathcal{V}_m$ are less than δ.

  We employ the nearest neighbor interpolation[1] to calculate privacy budgets in the road network. To be specific, we design a new function, as given in (3.25), to quantify the weights that a connection can copy from its one-hop neighbors. The weight of connection v_i^m, $m \geq 2$ depends on multiple $(m-1)$-hop connections that are adjacent to v_i^m. We first calculate all possible weights from the adjacent $(m-1)$-hop connections of v_i^m and then take the maximum one as $W_{v_i^m}$. The possible weight of v_i^m calculated from v_a^{m-1}, which is an $(m-1)$-hop connection and adjacent to v_i^m, i.e., $W_{v_i^m}^a$, is given by

$$W_{v_i^m}^a = \frac{\frac{D_{\max} - D_{v_i^m} + 2}{d_{\mathcal{G}}(v_a^{m-1}, v_i^m)} \times W_{v_a^{m-1}}}{\sum\limits_{v_j^m \in \mathcal{A}_a^m, j \neq i} \frac{1}{d_{\mathcal{G}}(v_a^{m-1}, v_j^m)} + \frac{D_{\max} - D_{v_i^m} + 2}{d_{\mathcal{G}}(v_a^{m-1}, v_i^m)}}, \tag{3.25}$$

  where $D_{v_i^m}$ is the degree of connection v_i^m (the sum of in-degree and out-degree); $D_{\max}$ is the maximum degree of connections in the segment; $\mathcal{A}_a^m$ is a set of m-hop connections that are adjacent to connection v_a^{m-1}. As road networks can be treated as strong-connected directed graphs, the degrees of the connections are at least two. $W_{v_i^m}$ can take the maximum one across different $W_{v_i^m}^a$, i.e., $W_{v_i^m} = \max(W_{v_i^m}^a)$.
- **Finalization (Steps 17-21):** The weights of connections which are less than δ are set to δ. The weight of a segment, i.e., W_{G_i}, takes the maximum weight of connections in the segment, and the privacy budget of the segment is calculated based on its weight, i.e., $\epsilon_{G_i} = \frac{\epsilon_0}{W_{G_i}}$.

There are two key parameters in (3.25), i.e., the degree of the connection $D_{v_i^m}$ and the shortest route distance to the last hop connections, $d_{\mathcal{G}}(v_a^{m-1}, v_i^m)$. Firstly, low-degree connections imply travel directions clearer than high-degree connections because low-degree connections offer fewer choices than high-degree connections. Thus, low-degree connections should be protected better

[1]The general nearest neighbor interpolation in image processing infers the pixel of a point by copying the pixels of its nearest neighbors.

Algorithm 2 Personalization Algorithm.

Input:
The road network, $\mathcal{G}$;
The weight threshold, δ;
The set of sensitive locations, $SL = \{sl_1, sl_2, \cdots, sl_n\}$;
The initial privacy budget, ϵ_0.

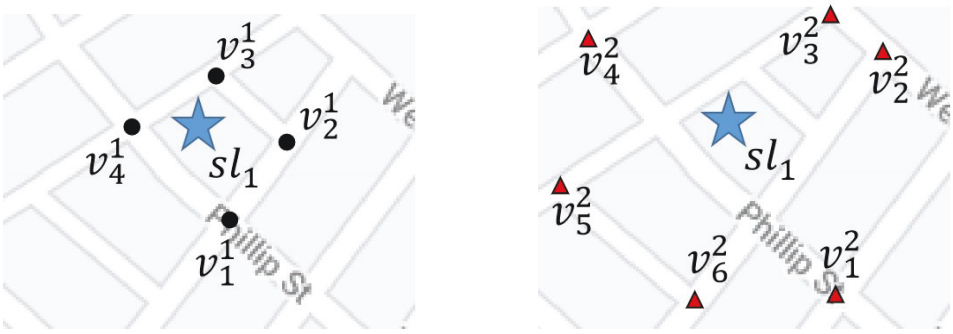
Output:
Privacy budget for each segment.

▷ **Initialization**
$\mathcal{V}_m, m = 1, 2, \cdots, M.$ ▷ Initialize connection sets
$W_{SL} \leftarrow 1$ ▷ Set the weights of sensitive locations
$W_v \leftarrow 1$, for all $v \in \mathcal{V}_1$
$W_v \leftarrow 0$, for all $v \in \mathcal{V}_i$, $i > 1$
▷ **Weight calculation**
$m = 2$
while $m \leq M$ **do**
    for $v \in \mathcal{V}_m$ **do**
        Calculate W_v ▷ With (3.25)
    end for
    if $\forall v \in \mathcal{V}_m$ has $W_v < \delta$ **then**
        break
    else
        $m++$
    end if
end while
$W_v \leftarrow \delta$, if $W_v < \delta$
▷ **Finalization**
for Segment G_i in road networks **do**
    $W_{G_i} \leftarrow \max_{v \text{ in } G_i} \{W_v\}$.
    Calculate $\epsilon_{G_i} = \frac{\epsilon_0}{W_{G_i}}$ ▷ The segment privacy budget
end for
return The privacy budget for each segment ϵ_{G_i}

than high-degree connections. Secondly, connections near sensitive locations are expected to be protected better than connections that are far away from sensitive locations.

The proposed personalization algorithm customizes privacy budgets for the connection-interval obfuscation according to drivers' sensitive locations and the features of road networks. In contrast, existing works (e.g., [73]) use a uniform private budget across a road network. The personalization algorithm in the PLPP scheme runs once when drivers configure sensitive locations and thus does not slow down the connection-interval obfuscation.

Given a road network and sensitive locations, the connection-interval obfuscation can use the private budget ϵ, the obfuscation radius r, and the number of intervals per road α, to perturb locations. The private budget ϵ is calculated using the proposed personalization algorithm with the initial obfuscation privacy budget ϵ_0 and the connection weight threshold δ. A driver can use a small ϵ_0 and/or a large r for large AEs and long shift distances. The driver can also



(a) The one-hop connection set $\mathcal{V}_1$.

(b) The two-hop connection set $\mathcal{V}_2$.

Figure 3.3: An example of $\mathcal{V}_1$ and $\mathcal{V}_2$, where sl_1 is a sensitive location.

configure a large α for fine interval perturbation.

3.5 Experimental Results

In this chapter, we evaluate the PLPP scheme with two real-world trajectory datasets, i.e., GeoLife GPS Trajectories (182 drivers with 17,621 trajectories in Beijing, China) [341] and T-Drive trajectory (10,357 divers in Beijing, China) [342].

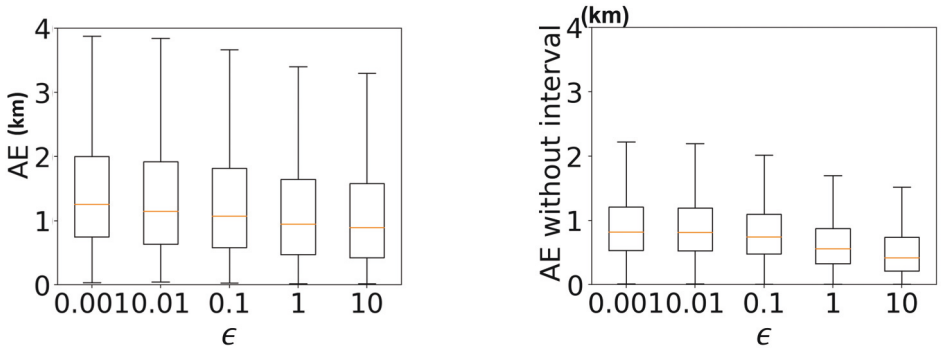
3.5.1 Location Privacy Protection

We first compare the AEs of the PLPP scheme and a connection-only scheme, i.e., [73]. We select the vehicle trajectories from the two datasets within the same region[2]. We build a simulation road network with clear connections in the region and then vertically project trajectories onto the nearest roads if they are not on the simulation road network. The road network is evenly divided into ten segments. The privacy budgets of the segments are calculated using Algorithm 2 with ten random sensitive locations on trajectories and the initial privacy budget ϵ.

The experimental results reveal that the PLPP scheme protects location privacy better than the connection-only scheme [73], as shown in Fig. 3.4. The boxplots indicate that the AEs of the PLPP scheme evenly distribute across longer ranges than the AEs of the connection-only scheme. The interquartile ranges in Fig. 3.4(a) are almost twice as large as those in Fig. 3.4(b) because adversaries cannot infer the obfuscated connections after the interval perturbation in the PLPP scheme. The medians of the AEs of the PLPP scheme are longer than those of the scheme in [73] across all ϵ. The median is about double under the PLPP scheme (1 km) than it is under the connection-only scheme (0.5 km) when $\epsilon = 10$. The medians in the two subfigures decrease with the increase of ϵ, indicating that the drivers can set small ϵ for high location privacy protection.

The indistinguishability of the PLPP scheme, i.e., $\frac{\Pr[v'|v_0]}{\Pr[v'|v_1]} \leq e^{\epsilon d_{\mathcal{G}}(v_0,v_1)}$ in (3.17), is validated in Fig. 3.5. In this experiment, the obfuscation radius is 400 m. We select two actual locations v_0 and v_1 when $d_{\mathcal{G}}(v_0, v_1) = 1$ km and run the PLPP scheme for the obfuscated locations v'. As shown in Fig. 3.5, the left-hand side of (3.17) is bounded, validating the indistinguishability of the PLPP scheme. An adversary can hardly distinguish actual locations that have the similar

[2]The longitude and latitude of the region are in ranges of [116.3564, 116.3740], and [39.907, 39.9241], respectively.



(a) AE of the PLPP scheme

(b) AE of the scheme in [73]

Figure 3.4: The comparison of the AE of the proposed PLPP scheme and the scheme in [73].

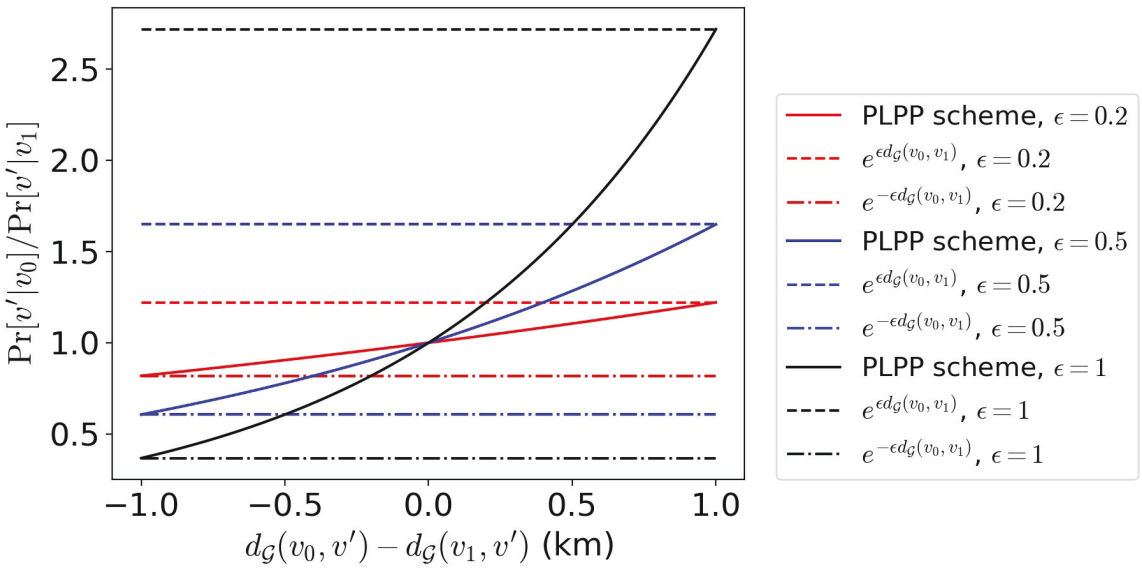


Figure 3.5: The indistinguishability of the PLPP scheme. The obfuscation radius is 400 m.

route distance from the obfuscated location, i.e., $\frac{\Pr[v'|v_0]}{\Pr[v'|v_1]} = 1$, with $d_{\mathcal{G}}(v_0, v') - d_{\mathcal{G}}(v_1, v') = 0$. We also see that the left-hand side of (3.17) is approximately linear under a small ϵ, e.g., $\epsilon = 0.2$. The reason is that when x is much smaller than one, $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} \approx 1 + x$.

The influence of the obfuscation radius on the shift distance and AE is shown in Fig. 3.6. The obfuscation radius ranges from 300 m to 700 m. The other parameters are the same as those in Fig. 3.5. All the connections are in the same segment. The initial privacy budgets are 0.001 and 1 to simulate a strict privacy-preserving case and a mild privacy-preserving case, respectively. The shift distance and AE grow with the obfuscation radius approximately linearly. As a result, drivers need to trade data utility for location privacy, but the AEs are much larger (almost double) than the shift distances. The growth is approximately linear because the route distance between actual locations and obfuscated location candidates linearly increases with the rising obfuscation radius in even road networks. The shift distance and AE increase slower under $\epsilon = 1$ than those under $\epsilon = 0.001$. This is because the obfuscated locations under big ϵ are closer to the actual locations than the obfuscated locations under smaller ϵ.

We proceed to compare the PLPP scheme with a popular 2D location obfuscation scheme, i.e.,

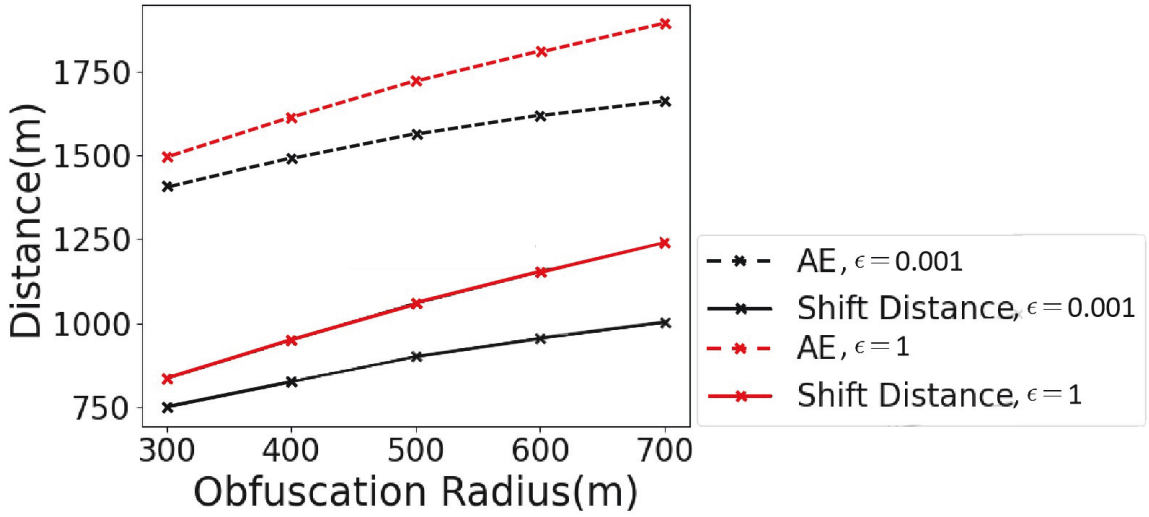
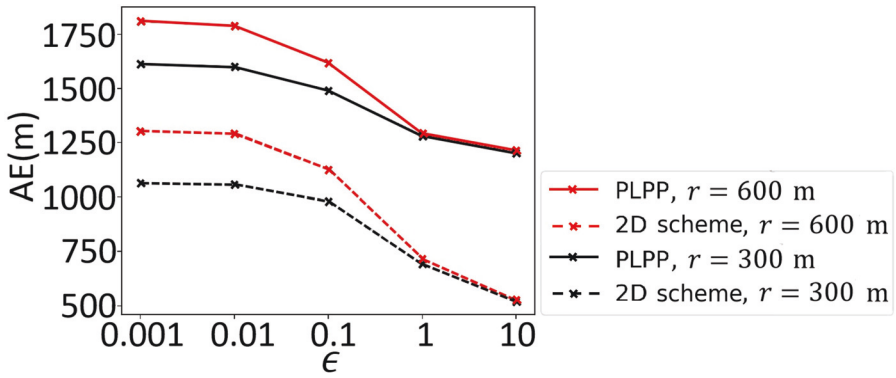
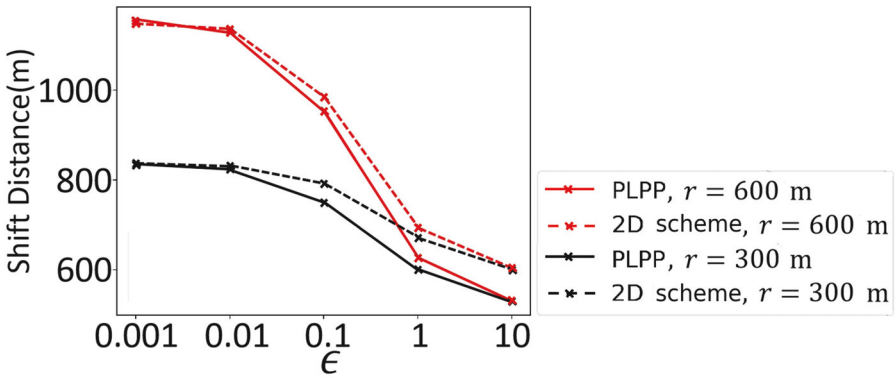


Figure 3.6: The influence of the obfuscation radius.



(a) Comparison of AE



(b) Comparison of shift distance

Figure 3.7: A comparison of the PLPP scheme and 2D Laplace scheme [9].

the planar Laplace scheme [77], [89]. The planar Laplace scheme treats a road network as a 2D plane and employs the Euclidean distance metrics. Fig. 3.7 compares the average AE and shift distance of the PLPP scheme and the planar Laplace scheme with GeoLife GPS Trajectories. For comparison purposes, all the segments use the same privacy budget ϵ in the PLPP scheme.

As shown in Fig. 3.7, the PLPP scheme can achieve a longer average AE and shorter average shift distance than the scheme in [77], especially in the case of high privacy budgets. The gap between the two schemes grows with the increasing privacy budget. The PLPP scheme can achieve a shorter average shift distance when ϵ is higher. The average shift distance measured by the route distance can be longer than the obfuscation radius because the route distance is no shorter than the Euclidean distance. The PLPP scheme has a similar average shift distance to the 2D Laplace scheme under $\epsilon = 0.001$. The reason is that locations are obfuscated with similar probabilities in the two schemes when the privacy budget is 0.001. When the privacy budget is 10, the average shift distance and the average AE with a 300 m radius are similar to those with a 600 m radius. This is because the obfuscated locations are close to actual locations in the case of low privacy requirements (high ϵ). The difference between shift distances under the two radiuses also decreases with the increasing privacy budget, as shown in Fig. 3.7(b). This is because both schemes are likely to select close locations as the obfuscation results under high privacy budgets. The shift distance of the PLPP scheme is shorter than the 2D planar Laplace scheme in road networks.

3.5.2 Personalization Algorithm

We evaluate the privacy budget of a connection depending on its degree and route distance to a sensitive location in the personalization (Algorithm 2) of the PLPP scheme. We consider a 1-hop connection v_1^1 and a 2-hop connection v_1^2. The connection v_1^1 has other adjacent connections v_i^2 $(i \geq 2)$. $D_{\max} = 8$, and initial privacy budget $\epsilon_0 = 0.1$. The route distance $d_{\mathcal{G}}(v_1^1, v_1^2)$ is from 50 m to 1,000 m. It is shown in Fig. 3.8 that the privacy budget of connection v_1^2 increases, as $d_{\mathcal{G}}(v_1^1, v_1^2)$ and $D_{v_1^2}$ increase. The ϵ of a connection in the PLPP scheme is lower, as compared to the scheme that does not consider the degree of connections (i.e., setting $(2 + D_{\max} - D_{v_i^m})$ in (3.25) to 1).

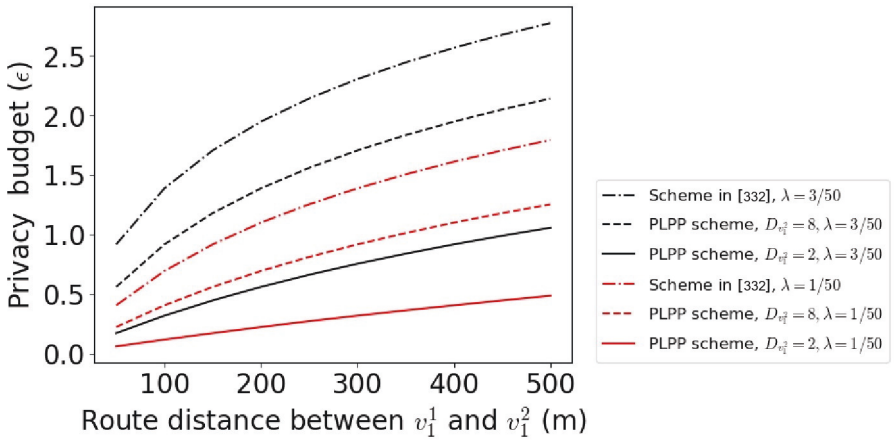


Figure 3.8: The impact of the degree of a single connection and route distance to a sensitive location in the personalization algorithm.

We then evaluate the privacy budget of multiple connections on a continuous road affected by the number of connections, the route distance to the first connection v_1^1, and the degree of connections. We consider a 1000 m road starting from a 1-hop connection on which five or eleven connections are evenly spaced, and evaluate the privacy budget of each connection. The m-th connection on the road is the first m-hop connection v_1^m $(1 \leq m \leq 5$ or $1 \leq m \leq 10)$. Other parameters are as follows, $D_{\max} = 10$, $\sum_{i \geq 2} \frac{1}{d_{\mathcal{G}}(v_1^j, v_i^{j+1})} = \frac{1}{100}$ $(1 \leq j \leq m-1)$, and

$\epsilon_0 = 0.1$. It is shown in Fig. 3.9 that, with the same route distance, the privacy budget of a connection increases when the number of connections on the road grows. The degree of a connection influences the privacy budget allocation. If a connection has a high degree, its privacy budget increases faster than its low-degree counterparts. Some schemes set a global privacy budget [343] or only consider the distance between two connections [344]. Compared to our personalization algorithm, the privacy budgets of the schemes developed in [344] and [343] increase faster, as shown in Fig. 3.9. Compared with the schemes (e.g., [343], [344]) which do not capture the connection degree, the proposed personalization algorithm ensures that a connection can inherit more privacy protection from the sensitive location.

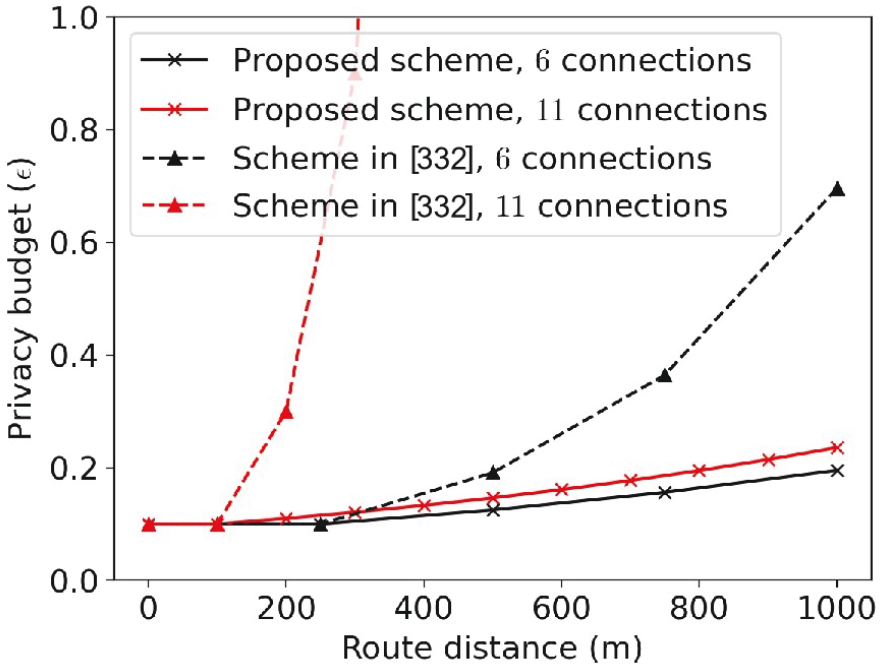


Figure 3.9: The privacy budget of connections on a road impacted by the number of connections, the route distance to the first connection v_1^1, and the degree of each connection.

A comparison study is carried out between the proposed personalization algorithm and the personalization algorithm developed in [89]. The scheme in [89] sets the privacy budgets according to the distance between the current location and the last inferred location. We randomly select 20 trajectories in the simulated road network, where each trajectory consists of 10 to 20 actual locations. The privacy budgets calculated with the two personalization algorithms are combined with the proposed connection-interval perturbation scheme. The obfuscation radiuses of 400 m and 600 m are used. As shown in Fig. 3.10, the proposed personalization algorithm achieves shorter shift distances (i.e., better data utility) than the existing algorithm developed in [89], especially for small ϵ_0. For example, the obfuscation with the proposed personalization algorithm provides the shift distance of 700 m when $\epsilon_0 = 0.1$ and $r = 400$ m. In contrast, the obfuscation with the algorithm developed in [89] achieves a shift distance of over 1,100 m. The shift distance of the PLPP scheme decreases faster than that of the existing algorithm, indicating that the proposed personalization algorithm can improve data utility. The shift distances of the two algorithms are similar, when the privacy budget is high, e.g., $\epsilon_0 = 10$. This is because the two schemes tend to select connections close to the actual locations to be obfuscated connections, when data utility is preferred over privacy.

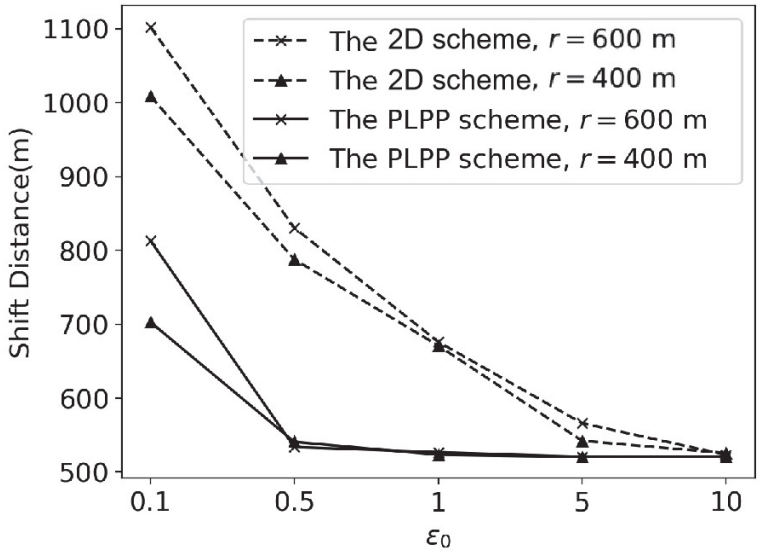


Figure 3.10: A comparison of the personalization algorithms in the PLPP scheme and in [89].

3.5.3 Influence of the Road Network

Table 3.3: Comparison of running time (ms).

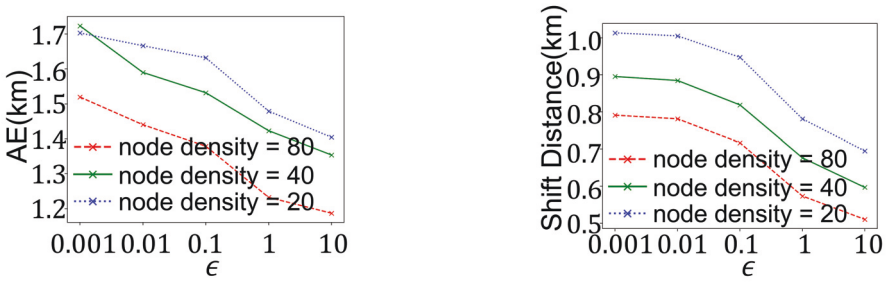
Road density (connections/km^2)	The PLPP scheme	The interval perturbation in the PLPP scheme	The personalization algorithm in [89]	The scheme in [73]
100	0.1954	0.0060	0.2275	0.2772
150	0.3091	0.0059	0.3670	0.5426
200	0.4029	0.0059	0.4947	0.7201
250	0.4907	0.0060	0.6083	0.8059

We generate virtual road networks with different node densities to show the impact of road networks on the PLPP scheme. The node (connection) densities are set to 80/km^2, 40/km^2, and 20/km^2. The trajectories are from the T-Drive dataset. We select 100 trajectories randomly and map the trajectories in the generated road networks.

An adversary has smaller AE in the high node density regions because locations estimated by the adversary are closer to the actual locations. With 80/km^2 node density, the average AE of the PLPP scheme is the shortest among the three given node densities. In Fig. 3.11(a), we see that the AE in the network with the node density of 40/km^2, when $\epsilon = 0.001$ is slightly higher than the AE in the network with the node density of 20/km^2. The reason is that the obfuscated region covers a limited number of obfuscation candidates when the node density is low. The limited connections may not provide the expected privacy-preserving capability.

The PLPP scheme can provide better data utility in high-density road networks than it does in low-density road networks, as shown in Fig. 3.11(b). The route distances between obfuscation candidates and actual locations in high-density networks are shorter than in low-density networks. Thus, the expected average shift distances in the high-density regions are shorter than those in the low-density regions. According to Figs. 3.11(a) and 3.11(b), the curves with different node densities show similar trends, as ϵ increases.

The impact of the interval number α on the interval selection probability is evaluated in Fig. 11.



(a) AE (b) Shift distance

Figure 3.11: The impact of the node density on AE and shift distance.

The number of intervals per road, i.e., α, can be configured using machine learning techniques, e.g., [345] and [346], and evolutionary computing algorithms, e.g., [347]. In this chapter, we configure α based on Taguchi's method [345], which reduces experimental tests by separating orthogonal parameters. In the PLPP scheme, only α and ϵ are used in the interval perturbation and orthogonal to other parameters, e.g., r. In Fig. 11, $\alpha \in \{15, 20\}$ and $\epsilon = 1$. The length of the road is 90 m, where the driver is 13 m away from the connection and its actual interval index is 3 for both $\alpha = 15$ and 20. The obfuscation probability is horizontally symmetric since the interval u_i is obfuscated in the same way as $u_{\alpha-i}$; see (7).

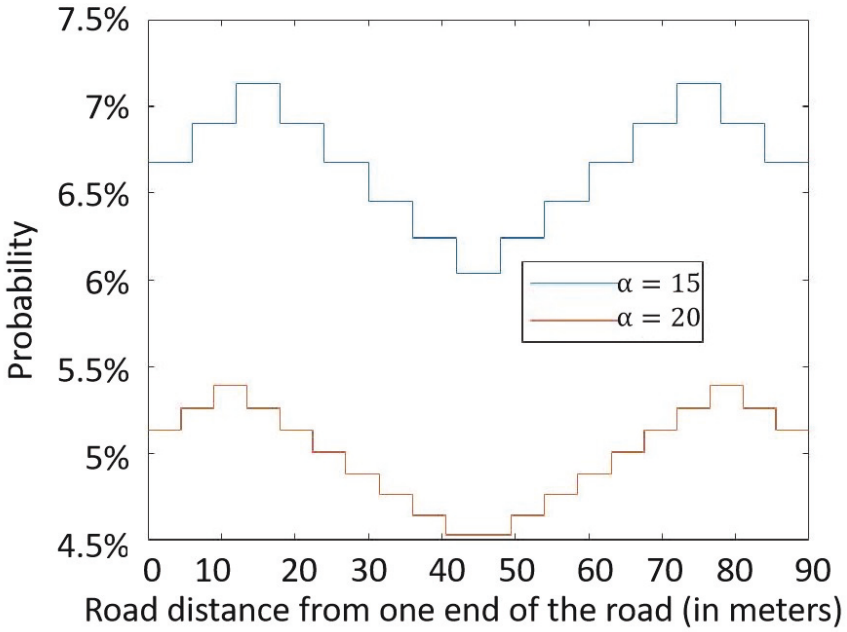


Figure 3.12: The impact of the interval number α on interval selection, where α is set to 15 and 20. The privacy budget ϵ is set to 1.

3.5.4 Running Time

The PLPP is designed to minimize the latency due to obfuscation by enabling privacy budgets, the connection perturbation distribution, and the interval perturbation distribution to be pre-calculated. Before any LBS requests, the PLPP scheme can calculate privacy budgets for all connections and then build constant connection perturbation and interval perturbation tables.

Each entry of the tables gives the perturbation probability of a pair of connections (3.6) or intervals (3.7). The sizes of the connection perturbation and interval perturbation tables is $O(n^2)$ and $O(\alpha^2)$, respectively, where n is the number of connections in the road network. During the LBS requests, the connection-interval obfuscation can look up the tables to probabilistically select the obfuscated connections and locations.

We compare the running time of the proposed PLPP scheme with the personalization algorithm in [89] and the non-personalized obfuscation algorithm in [73], as shown in Table III. The personalization algorithm developed in [89] sets privacy budgets based on the distance between the current and the last inferred locations. The obfuscation algorithm presented in [73] employs a 2D single-obfuscation and maps the obfuscated locations to the nearest connections in the road network. To show the running time of each module, we combine our dual-obfuscation with the personalization algorithm in [89] and our personalization algorithm with the 2D single-obfuscation [73]. The other parameters are as follows: 500 actual locations are randomly selected, the obfuscation radius r is 500 m, and the initial privacy budget ϵ_0 is 1.

As shown in Table III, the proposed PLPP scheme is faster than the dual-obfuscation with the personalization algorithm in [89] in all cases with up to 19% improvement (i.e., 0.4907 ms versus 0.6083 ms). The proposed personalization algorithm only runs once to calculate the privacy budgets for all connections when drivers configure their sensitive locations before obfuscation. In contrast, the personalization algorithm in [89] needs to be executed for each obfuscation. The PLPP scheme is also faster than the 2D single-obfuscation [73]. This is because the 2D single-obfuscation [73] needs to map the obfuscated locations onto the road network, which is time-consuming. Table III also shows that the running time of the interval perturbation is only about 0.006 ms. Compared to that of the whole PLPP scheme, the overhead of the interval perturbation is negligible. Moreover, the running time of the interval perturbation does not grow with the connection density. The reason is that the small constant α (i.e., 5 in the experiment) and the same interval perturbation distribution are applied to all roads.

3.6 Conclusion

In this chapter, we propose the RN-I to evaluate obfuscation-based location privacy-preserving schemes in road networks. We propose the PLPP scheme to protect the location privacy of vehicles. The PLPP scheme employs a dual location obfuscation consisting of a connection perturbation and an interval perturbation. A location is first mapped to a remote connection in the connection perturbation and then pseudo-randomized around the connection in the interval perturbation. A personalization algorithm is developed to customize the privacy budgets of connections based on the sensitive locations specified by drivers. The proposed PLPP scheme is proven to achieve RN-I and validated with comprehensive experiments using two real-world trajectory datasets.

In the proposed scheme, a single vehicle is considered regarding its location privacy. The route distance between any two mutually reachable locations is utilized to evaluate the indistinguishability of the locations for the vehicle. In our future work, we will generalize the proposed scheme to capture other important road network features, such as velocity, acceleration, heading, and surrounding traffic condition. We will also consider potential collaboration among multiple nearby vehicles to preserve their location privacy cooperatively.

Chapter 4

Enhanced Privacy Protection

In this chapter, we first propose a Convolutional Neural Network (CNN) based detection mechanism to detect illegal trajectories without requiring the drivers' actual location. The proposed scheme has high accuracy in detecting illegal trajectories even if the drivers protect actual location data with various noise sizes. By using the RN-Indistinguishability, we then design the Cloaking Region Obfuscation (CRO) mechanism that employs the route distances to quantify the indistinguishability of locations on roads. We prove that the CRO mechanism satisfies the RN-Indistinguishability. The CRO mechanism can be extended with general road network features without breaching differential privacy.

The rest of this chapter is organized as follows.

- Chapter 4.1 and Chapter 4.2 study the background of vehicular trajectory protection and detection and the existing works of vehicle trajectory obfuscation and detection, respectively. Chapter 4.3 describes the proposed trajectory detection mechanisms. Chapter 4.4 evaluates the proposed trajectory detection mechanisms with real-world road network datasets. Chapter 4.5 concludes the proposed vehicular trajectory protection and detection mechanism.
- Chapter 4.6 provides the background of location and identity privacy protection in vehicular networks. Chapter 4.7 studies the existing works of location and identity protection. Chapter 4.8 illustrates the system and adversary model. Chapter 4.9 shows the proposed CRO mechanism, which is evaluated in Chapter 4.10. Chapter 4.11 gives a conclusion of the proposed CRO mechanism.

4.1 Introduction of Vehicle Trajectory Obfuscation and Detection

Location-based services (LBS) have been extensively and deeply developed as an important part of smart cities to provide various services, e.g., traffic analysing and urban planning [348], [349]. Smart cities inevitably require spatio-temporal location data of vehicles, which is highly correlated with a driver's private information (e.g., home address, company address, and religion) [350]. LBS providers rely on location data from drivers to offer various services [351]. However, adversaries can infer drivers' personal information by analyzing location data [352]. It is of necessity to protect location data in Internet-of-vehicles (IoV) for drivers' privacy while

ensuring high quality of service (QoS).

Obfuscation schemes [211] and adaptive location privacy-preserving schemes [89] have been developed to protect location privacy. Obfuscation schemes add noise to obfuscate drivers' actual location data, which reduces the data utility of the location data. The obfuscated location data are indistinguishable from each other, which leads to the fact that LBS providers would collect illegal location data. The location privacy-preserving schemes aim to balance the QoS of LBS and privacy-preserving capability by analyzing the drivers' requirements.

Malicious drivers breach legal drivers' profit. For example, malicious drivers can occupy more benefits than they deserve by deliberately modifying their trajectory data in Taxi service [84], [85]. Malicious drivers also use location privacy-preserving schemes to protect their location data. By analyzing location data which concludes illegal location data, smart city applications cannot provide an acceptable QoS of LBS. Therefore, LBS should detect the illegal location data to ensure high QoS. If the malicious drivers employ location privacy-preserving schemes (e.g., obfuscation schemes) as the legal drivers, detecting illegal data becomes difficult.

We study the illegal location data detection and propose a personalized obfuscation scheme and an illegal trajectory detection mechanism. Our work has two-fold contributions as follows:

- We employ the differential privacy in the proposed personalized obfuscation scheme to protect drivers' location privacy adaptively and to provide high QoS of LBS in road networks.
- We propose a CNN-based detection mechanism to detect illegal trajectories without requiring the drivers' actual location. The proposed scheme has high accuracy in detecting illegal trajectories even if the drivers protect actual location data with various noise sizes.

We conduct experiments with the real-world road network dataset extracted from Open Street Map (OSM)[1] to evaluate the proposed scheme.

The illegal trajectories are generated with fake speeds and paths with the real-world road network dataset. We also evaluate the proposed scheme in the case that the drivers adaptively protect their location data, i.e., using different privacy levels. The experimental results show that the proposed detection scheme achieves at least 94% Area Under the Curve (AUC) score when detects the illegal obfuscated trajectory.

4.2 Related Works of Vehicle Trajectory Obfuscation and Detection

The existing smart city provides location-based services by mining trajectory data that are transmitted in vehicular networks [353]. Wang et al. have studied the privacy challenges in smart city and analyzed the privacy leakage in LBS [323]. The authors pointed out that the smart city can provide high quality of services if trajectory privacy is well protected.

The previous studies of location obfuscation mechanisms perturb a driver's actual location and report an obfuscated version to LBS. Derived from the differential privacy [354], scheme in [211] first developed the concept of geo-indistinguishability. The scheme follows the idea of geo-indistinguishability and uses Laplace distribution to add controlled noise for protecting

[1]Open Street Map is an open source database of the world's geographic map. `https://www.openstreetmap.org/`

location data locally. Yu et al. [355] improved the two-phase dynamic differential location privacy scheme by integrating the inference error expectation and geo-indistinguishability [356]. The improved framework effectively protects location privacy in a 2D map. The authors developed an adaptive location privacy-preserving mechanism in [89] to balance location privacy and utility. The mechanism calculates the amount of noise before adding noise to actual location data. The calculation is based on the correlation level between the driver's current location and the previous obfuscated locations. With the concept of differential privacy, Xiao et al. [357] improved a location-cloaking system to protect drivers' location data in a 2D map. The obfuscation locations generated by the existing 2D obfuscation mechanisms might locate at unreachable locations, e.g., in the river, which breaches the location privacy-preserving capability of the obfuscation mechanisms.

The existing illegal trajectory detection mechanisms are classified into the machine-learning-based detection and the rule-based detection. The illegal trajectory data is detected by utilizing GPS data in the rule-based detection mechanisms. Machine-learning-based detection mechanisms classify trajectory data as legal and illegal by using techniques like deep neural networks. In [358], Chen et al. improved an efficient real-time trajectory detection method with low processing overhead. The method uses the window size to estimate the partial trajectory that result in the anomalousness trajectory. The trajectory detection system with two-phase outliers upon trajectory data streams was improved in [359]. The two phases are the trajectory simplification and the outlier detection.

In [360], authors developed a trajectory detection method based on a recurrent neural network (RNN). The authors extracted drivers' behaviors within a sliding window and uses the deep representations that are fixed-length for the feature sequence. The authors grouped the representations into clusters before detection. CNN is first introduced in [361], which is further utilized in the fields such as the natural language processing and speech recognition. A CNN-based trajectory prediction method was improved in [362]. The method simplifies the network structure and utilizes the trajectory structure (spatio-temporal consistency). The experimental results show that the CNN-based trajectory prediction method can detect illegal trajectories with a high score of the AUC.

Our work obfuscates drivers' trajectories in the road networks to avoid the generation of the off-road obfuscated locations. Then, an illegal trajectory detection scheme based on CNN is proposed in this chapter. The proposed scheme does not expose the drivers' actual trajectories and achieves high detection accuracy. The proposed scheme can detect illegal trajectories even if the drivers use various privacy parameters to obfuscate the locations. To the best of our knowledge, we are the first work that detects illegal locations in real road networks, which is almost the actual usage scenario of a smart city.

4.3 Proposed Scheme of Vehicle Trajectory Obfuscation and Detection

In this chapter, we start by proposing an adaptive obfuscation scheme to customized protect location privacy in road networks. Then, we propose an illegal trajectory detection system with CNN to identify the legal and illegal trajectory from the obfuscated trajectory. In our model, X_m is the m-th trajectory consists of the location points sequence $(x_{m1}, x_{m2}, \ldots, x_{mn})$. $x_{mi} = (lat_i, lon_i)$ is a tuple which stands for the coordinate (i.e., latitude lat_i and longitude lon_i) of a location.

4.3.1 System Model

The existing obfuscation studies [89], [363], [364] pay attention to protect the drivers' location data in a 2D map, which generate off-road locations (e.g., railroads and rivers). Adversaries can exclude the off-road obfuscated locations from real trajectory data when they identify the obfuscated locations that are off-road. The Euclidean distance between the obtained off-road location and the nearby road can be utilized by the adversaries to estimate the actual location. In this chapter, we controlled the obfuscated candidates to avoid the off-road data and guarantee that all obfuscated locations are on-road.

4.3.2 Dynamic Obfuscation Scheme

Definition 6. ***Geo-indistinguishability*** *[211]: Let P be a probabilistic function. Let X and Z be a set of the actual location candidates and obfuscated locations candidates, respectively. The K represents the mechanism that uses the probability $P(Z)$ to map an element in X to an element in Z. K is ϵ-geo-distinguishable, if and only if for all x, x' has:*

$$d_{\mathcal{P}}\left(K(x), K\left(x'\right)\right) \leq \epsilon d\left(x, x'\right) \tag{4.1}$$

The $X_m = \{x_{m1}, \ldots, x_{mn}\}$ is the raw path that indicates the actual trajectories, while $Z_m = \{z_{m1}, \ldots, z_{mn}\}$ indicates the obfuscated trajectories. We utilize 2D Laplace noise $D_\epsilon(x)(z) = \frac{\epsilon^2}{2\pi} e^{-\epsilon d(x,z)}$ as obfuscation distribution in this chapter. The reason is that the 2D Laplace noise ensures that z_{mi} is distributed around x_{mi}. The probability of z_{mi} with 2D Laplace noise decreases with the increasing of $d(x_{mi}, z_{mi})$ ($d(\cdot, \cdot)$ is the Euclidean distance in this chapter). Theϵ-geo-distinguishable privacy condition is also satisfied with the 2D Laplace noise.

4.3.3 Adaptive location privacy-preserving scheme

A new adaptive location privacy-preserving scheme is proposed in this chapter. The proposed scheme sets ϵ correlated to the generated obfuscated locations. By utilizing the proposed adaptive location privacy-preserving scheme, we boost the randomness of the noise generation, which increases the difficulty of inferring the actual location of a driver.

We configure ϵ into the high, medium, and low privacy levels. We set the average distance of the obfuscated location to r when there is a low level ϵ. A medium and a high level ϵ have an average distance of $1.5r$ and $2.25r$, respectively. The proposed scheme obfuscates every single location point in each continuous trajectory. As the starting location and destination are more sensitive to a driver, we set the highest privacy level for the two locations. For other locations x_i in the trajectory, the obfuscation parameters are related to the Euclidean distance $d(x_i, z_{i-1})$. We set two different thresholds $D1$ and $D2$ to divide $d(x_i, z_{i-1})$ into three types, where $D1 \neq D2$. When the value of $d(x_i, z_{i-1})$ is bigger than the values of $D1$ and $D2$, the correlation between x_i and z_{i-1} is weak. In this case, a low-level noise is added in the actual location data. When the value of $d(x_i, z_{i-1})$ is less than the values of $D1$ and $D2$, x_i and z_{i-1} is close in road networks (i.e., high correlation). Therefore, we add noise with a high privacy level when obfuscating actual locations in this scenario. Otherwise, we set ϵ as a medium privacy level in the proposed obfuscation scheme.

The proposed scheme reduces the correlation between x_i and z_{i-1}. Hence, the adversary cannot infer ϵ by analyzing the prior knowledge and the obtained obfuscated locations within a specific

time window. The adversary cannot predict the driver's future locations because the value of ϵ is changing, which increases the difficulty of attacks.

The selection of ϵ in the proposed scheme aims to balance location privacy and data utility. A large amount of noise is required to achieve a high privacy level but leads to a low QoS of LBS. The different ϵ can provide customized geographic location accuracy which suit for various LBS requirements. For example, location-sensitive LBS (e.g., navigation) needs a high accuracy location data so that the scheme ought to utilize a high ϵ to provide a high data utility. For location-insensitive LBS, e.g., weather forecasts, the scheme can employ a low ϵ for a high privacy level.

The amount of the added noise is controlled in the proposed scheme to balance data utility and location privacy. we use z within a region based on x to set the upper-bound of the QoS loss and that of the capability of location privacy, i.e., $d_{max}(x, z)$. If the distance $d(x, z)$ between the actual location x and the obfuscated location z exceeds $d_{max}(x, z)$, the proposed scheme will obfuscate the driver's actual location again.

The proposed scheme generates obfuscated locations $z_{m1}, \ldots, z_n$ and maps the obfuscated locations to the nearest road. Therefore, the proposed scheme obfuscates actual trajectories to reachable on-road locations.

4.3.4 Illegal trajectories detection based on CNN

A two-dimensional convolutional neural network (2D-CNN) model is applied in the proposed scheme to detect illegal trajectories. The proposed model is shown in Fig. 4.1.

One-dimensional CNN (1D-CNN) consists of the convolutional layer, the subsampling layer, and the optional fully-connected layers. In convolutional neural networks, the convolutional layer is the major part that analyzes the input data to extract classification features. The Separate feature extractor contains multiple convolution kernels. The convolutional layer of the CNN model can extract the Spatial-temporal correlation of the trajectory. After extracting features in convolutional layer. pooling step starts. The weight parameter redundancy is solved by the local connection and weight sharing. However, the over-fitting problems arise due to the CNN model degrades in the generalization performance. With the extracted features from the convolution layer and pooling, CNN model can reduce the data dimension while retaining the value of the principal feature map.

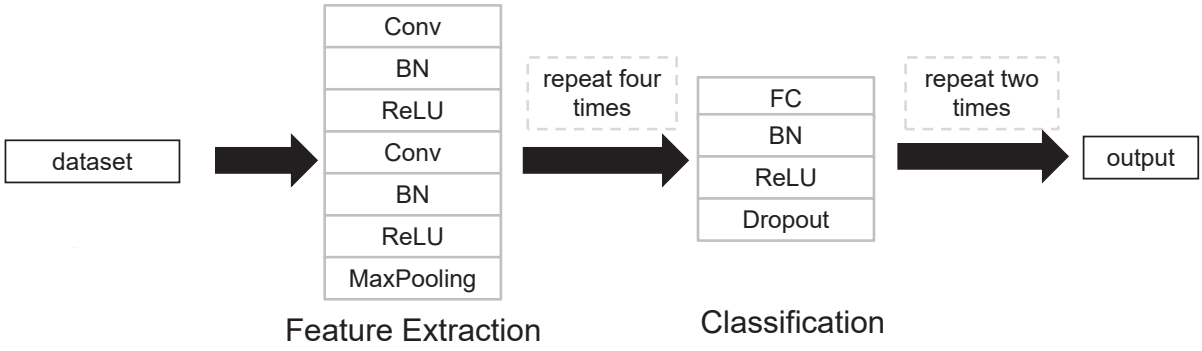


Figure 4.1: Our CNN model architecture.

Our work improves the architecture of 1D-CNN with 2D-CNN model, which is widely utilized in classification of images (e.g., ResNet [365], VGG [366], and GoogleNet [367]). The proposed model is described as follows.

- In advance of the maximum pooling layer, two convolutional layers are employed in our model. Thus, the proposed model can extract features effectively.
- We add a normalization (BN) layer after the two convolutional layers to retain the spatial-temporal correlation of locations. The fully connected (FC) layers are combined with the dropout layers and BN in the proposed model to avoid the FC layers leading to over-fitting issue.

4.4 Evaluation of Vehicle Trajectory Obfuscation and Detection

4.4.1 Original Dataset

We employ real-world road network information of Porto, which is extracted from OpenStreetMap (OSM) to evaluate the proposed scheme. OSM is popular in LBS applications, such as the route planning and the geocoding of address [368]. The extracted road network is shown in Fig. 4.2. The Portugal taxi trajectory dataset[2], whose recorded location data has 15 seconds time interval to build the trajectory, is used in our experiment.

4.4.2 Illegal trajectories

As far as we know, no public dataset are labeled with illegal and legal trajectories. Three methods are popular to generate illegal trajectories according to legal trajectories.

1. Insertion trajectory from other sources of trajectory dataset as the illegal trajectories [369].
2. Division the trajectories data into legal and illegal dataset [370].
3. Combination the above two methods [363].

In this chapter, we generate illegal trajectories by utilizing legal trajectories. The generated illegal trajectories are employed in our classification experiments. Hence, legal and illegal trajectories in the training and detection come from the same dataset to reduce deviation.

4.4.3 Simulation results

We evaluate the detection capability of the proposed scheme with the generated trajectory data and the public trajectory dataset of the Portugal taxi. In the experiment, we configure ϵ to control the level of Laplace noise.

Experimental setting

We use the public real-world trajectory dataset and the synthetic data as follows.

[2]Portugal taxi trajectory dataset[Online]. Available: https://www.kaggle.com/c/pkdd-15-predict-taxi-service-trajectory-i

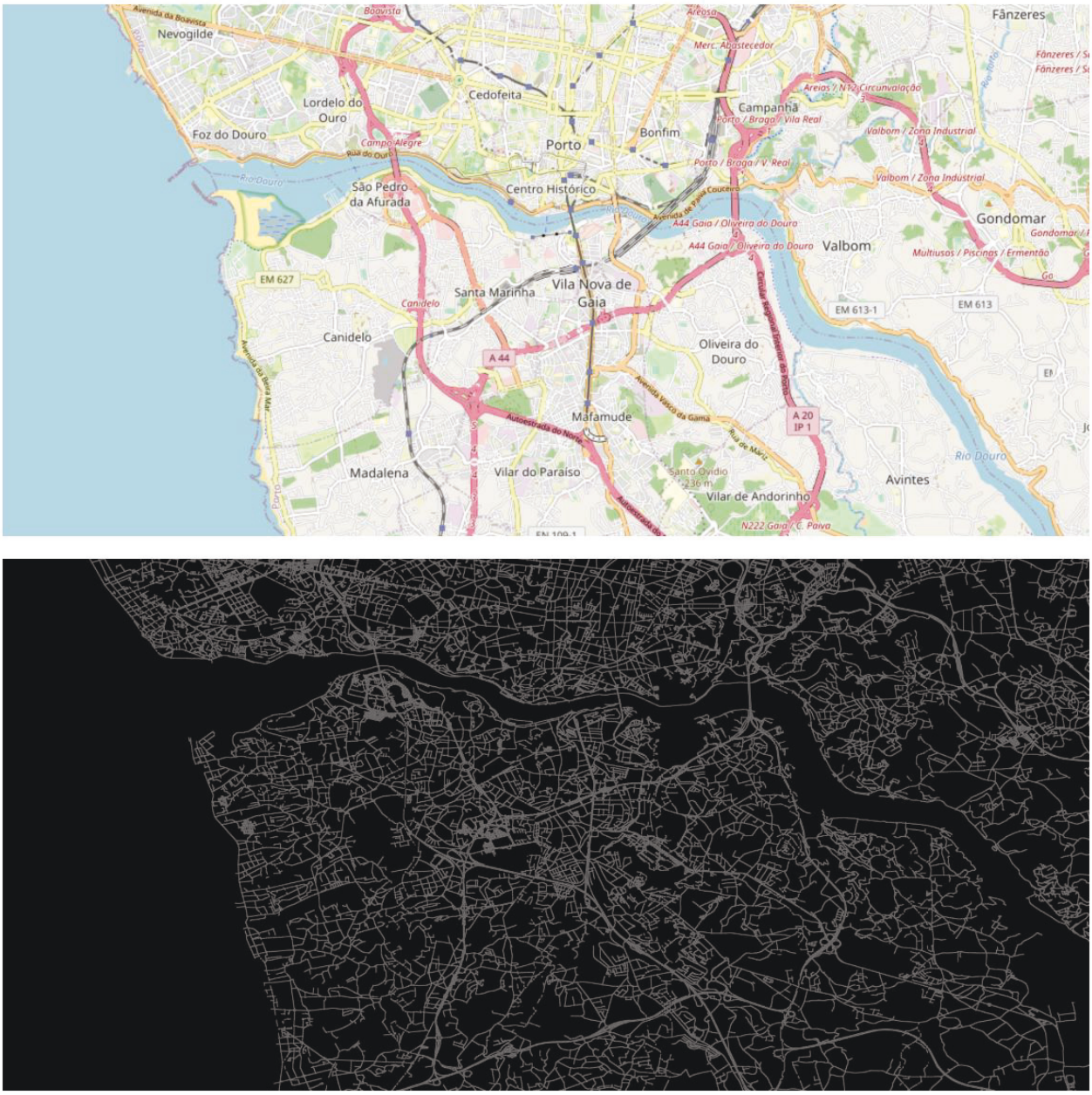


Figure 4.2: Generation of the road network. Upper: real road map in OSM; Bottom: the generated road network.

Real-world trajectory: We employ the public trajectory dataset of Portugal taxi which has approximate 1.7 million trajectory data.

Synthetic data: We extract half of the trajectory data from the Portugal taxi dataset to generate the illegal data. In this chapter, the illegal trajectories considered have two forms, speed anomaly and path anomaly. The two forms represent malicious actions, i.e., speeding and detour, respectively. The illegal trajectories are generated as follows:

- **Speed anomaly**: The maximum speeds of vehicles are limited in road networks. Malicious drivers drive faster than the limitations to obtain more profit within a period. The utilized trajectory dataset contains timestamp, we delete x_i in the selected continuous trajectory X with a certain probability and reassign the timestamp. The the trajectory X has a higher speed than the speed limitation.

- **Path anomaly**: Malicious drivers can also select a longer route than the recommended route, i.e., path anomaly. We utilize multiple legal trajectories (e.g., three trajectories X_a, X_b, X_c) to generate path anomaly illegal trajectories. The starting location and destination of X_a are denoted as x_{a1} and x_{an}, respectively, at shown in Fig. 4.3. We employ the trajectories X_b and X_c to intersect[3] the trajectory X_a at location x_{ai}, x_{aj}. We also combine the trajectories X_b and X_c at location x_{bl}. Then, we obtain an illegal trajectory which is a path anomaly. The starting location and destination of the generated trajectory are x_{a1} and x_{an}, respectively, but the route distance between x_{a1} and x_{an} is longer than it should be. In this chapter, the length of the generated trajectories are set to be at least 1.6 times as long as that of the legal trajectories.

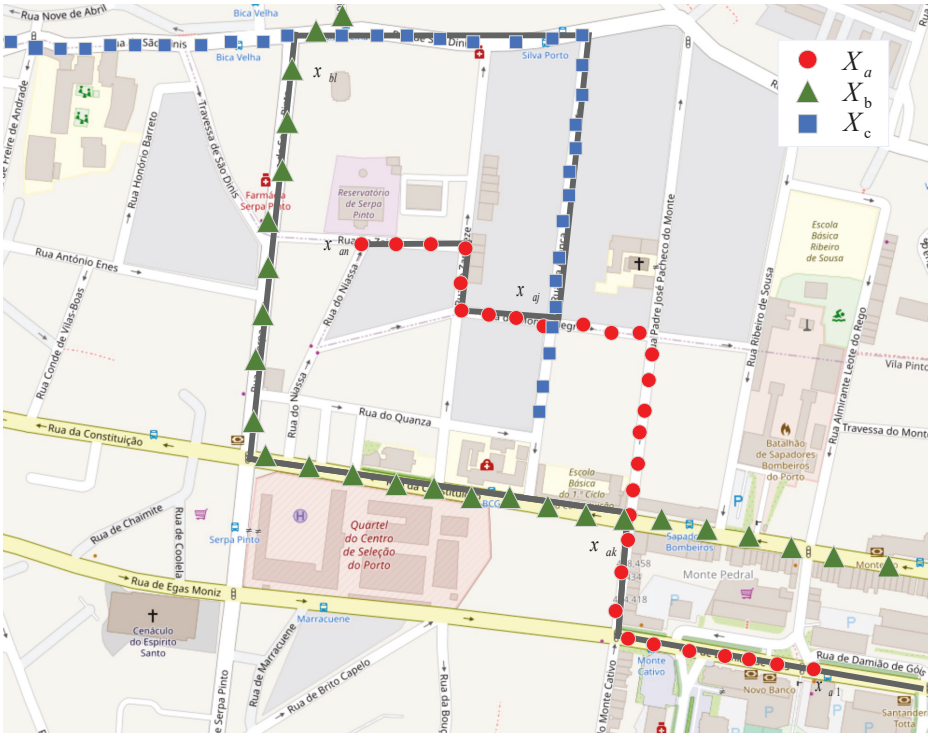


Figure 4.3: Example of an anomaly generated path.

[3]The intersection stands for the points of the two trajectories whose distance are within a certain range.

The above types of illegal trajectories include the most categories of malicious activities in the road networks. We use more than 600,000 legal and illegal trajectories to evaluate our proposed detection scheme.

Implementation

The configurations of the proposed mechanism that implemented in our experiments are as follows.

We use $\epsilon_0 = 0$ to stand for the non-protected situation. We configure the obfuscation radius as the average Euclidean distance between x_i and z_i when using noise level ϵ_i. When $\epsilon_0 = 0$, the average obfuscation radius is 0 m. We set ϵ_1 and ϵ_2 with obfuscation radii 100 m and 1000 m, respectively.

Experimental results

We use Python to conduct the experiments. We take the average value of the experiment results after running the experiment for five times.

Table 4.1: Experimental Accuracy

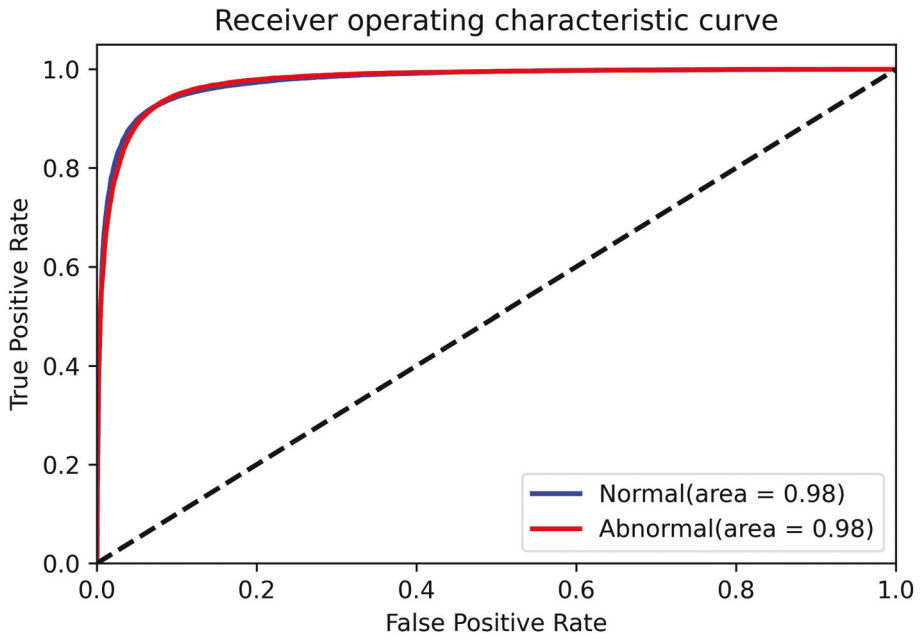
ϵ	**Average value of the noise radius**	**Accuracy**
ϵ_0	0 m	93.1%
ϵ_1	100 m	86.1%
ϵ_2	1,000 m	72.5%

A large amount of noise decreases the accuracy rate of the CNN model, as shown in Table. 4.1. We compare the receiver operating characteristics curve (ROC) with various setting of ϵ, as shown in Fig. 4.4. When applying ϵ_1 and ϵ_2, the proposed scheme achieves a high-value accuracy rate of 0.94 on the AUC score. Compared with that of ϵ_1, the AUC score of ϵ_2 has been reduced by 0.14 to 0.80.

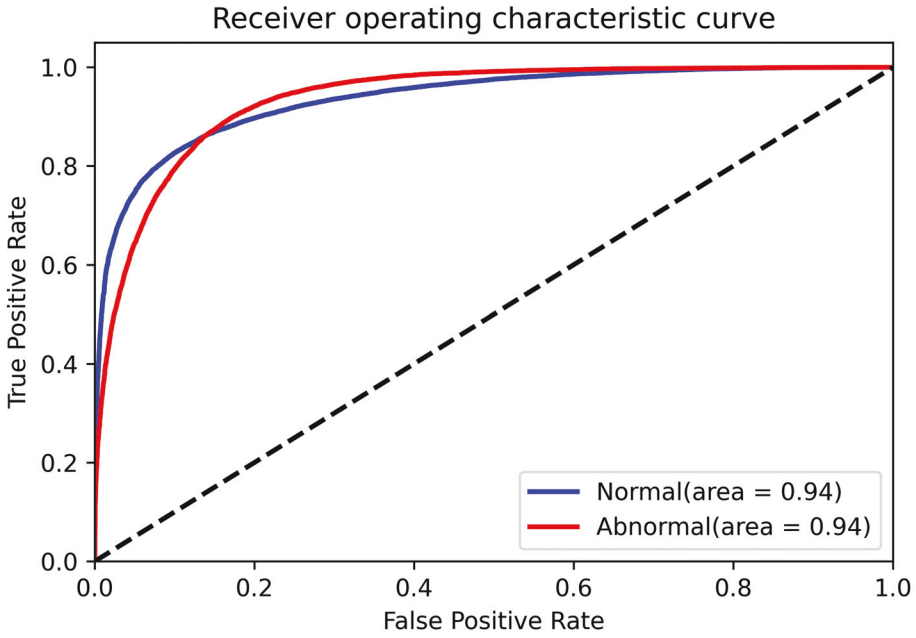
4.4.4 Contrast

We use the same trajectory dataset as the existing related work [363]. When detecting illegal trajectories, the proposed scheme achieves higher accuracy rate than the scheme developed in [363]. The proposed scheme considers a more complex environment, i.e., road networks, than the 2D plane environment considered in [363] The proposed scheme utilizes a different obfuscation process and recognition dataset, so we do not compare the accuracy of the two schemes. Compared with the schemes in [363], the proposed scheme has advantages as follows.

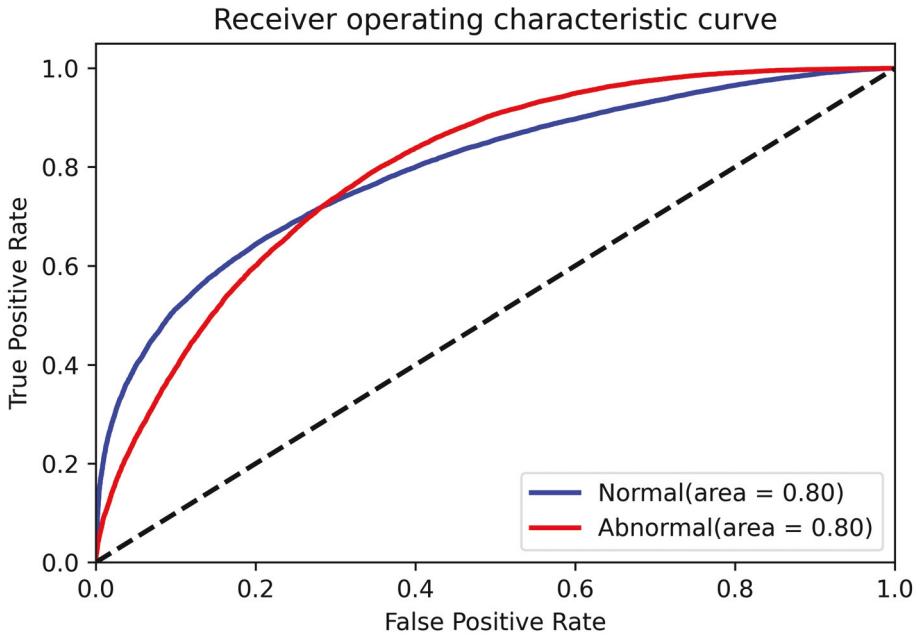
- The scheme in [363] employs fixed parameters in obfuscation process, while the proposed scheme dynamically calculate the parameters. The dynamically calculated parameters provide higher privacy protection capability and high data utility than the scheme in [363].
- The scheme in [363] manually inject trajectories to generate illegal trajectories. The illegal trajectories trajectories generated in this chapter are closer to the real world than that of the scheme in [363]. The generated illegal trajectories in this chapter are indistinguishable from the real trajectories which increases the difficulty to detect illegal trajectories. Under the strict assumption, the proposed scheme still achieve a higher accuracy rater than the existing work [363].



(a) ϵ takes 0 m, i.e., ϵ equals to ϵ_0



(b) ϵ takes 100 m, i.e., ϵ equals to ϵ_1



(c) ϵ takes 1,000 m, i.e., ϵ equals to ϵ_2

Figure 4.4: ROC under three different parameters [371].

- We employ the real-world dataset and road networks to evaluate the proposed scheme. Thus, the proposed scheme in this chapter has practical meaning.

4.5 Conclusions Vehicle Trajectory Obfuscation and Detection

In this chapter, we first propose a new scheme to adaptively protect location data in real-world road networks, which is an important scenario of smart cities. Then, we proposed an illegal trajectory detection scheme to detect illegal locations in the case that all drivers are protected in road networks. The privacy parameters of the proposed scheme was calculated by considering the correlation of the actual location and the obfuscated location. Thus, the adversary cannot infer utilized privacy parameters and the actual locations. We generated illegal trajectory data with speed anomaly and path anomaly to simulate the real-wold malicious driving. The 1D-CNN model with 2D-CNN architecture is proposed in detecting illegal trajectories. According to our experiment results, the proposed detection scheme achieve better performance (e.g., the AUC score is above 0.94) than the existing works in road networks.

In the future work, we will balance the data utility and location privacy to maximize the data availability while satisfying the requirements of drivers' privacy. Moreover, we will assess the privacy levels of driver's privacy and develop a new system to protect location data privacy with the capability to handle most drivers' requirements.

4.6 Introduction Cloaking Region Obfuscation for Road Network-Indistinguishability and Location Privacy

Location-based services (LBS) provide vehicular applications access with the location data [222]. LBS servers are able to obtain drivers' geographical locations, which puts drivers' location privacy at risk [372]. Adversaries can infer drivers' information from the location data after gaining unlawful access to the LBS servers [373]. Adversaries can also attack and eavesdrop on the communications between the LBS servers and drivers to obtain drivers' location data [8], [374]. Unlike the general mobile LBS users, vehicles are easier to be tracked due to the fact that the vehicles' trajectories are considerably predictable in road networks [375].

Obfuscation-based mechanisms have been developed to protect the LBS users' location privacy against eavesdroppers and malicious LBS servers [72]–[74]. Users can perturb their real locations locally by exploiting differential privacy mechanisms that add controllable noises to the location data [376]. The magnitudes of the additive noises under an indistinguishability constraint are determined based on the distance between an actual location and its obfuscated version [74]. Drivers can use the obfuscated location in their requests to an LBS server [69]. Although the obfuscation mechanisms can penalize location accuracy, they can still be used in location-insensitive LBS, such as location-based recommendations [76].

Vehicles in road networks have not been well considered in existing obfuscation-based mechanisms. General-purpose obfuscation-based mechanisms (e.g., two-dimensional (2D) Laplace location privacy-preserving mechanisms) use the Euclidean distance to measure the distance between an actual location and the corresponding obfuscated location [73], [74], [77]. However, the mechanisms underestimate the distance in road networks due to the fact that the Euclidean distance is less than, or equal to, the route distance between two locations [78]. Another limitation of the existing obfuscation-based mechanisms is that the mechanisms may generate off-road locations, e.g., locations in a river [79]. Extra effort (at a cost of high computational complexity) is needed to avoid those off-road locations [80]. In existing location obfuscation mechanisms [73], [74], [77], vehicles independently obfuscate their locations, which increases the risk of location data leakage [330].

In this chapter, we define a new notion of Road Network-Indistinguishability (RN-Indistinguishability), where the indistinguishability between locations is measured with road network features (e.g., road network density, speed limits, and route distance). The difference between the proposed RN-Indistinguishability and the existing Geo-I is shown in Fig. 4.5. Based on the RN-Indistinguishability, we propose a new Cloaking Region Obfuscation (CRO) mechanism that protects location privacy in road networks by considering road network features (e.g., route distance). Vehicles are obfuscated to other regions, according to the indistinguishability between the obfuscated region and the actual cloaking region. The CRO allows multiple vehicles to cooperate. Vehicles in the same cloaking region are uniformly obfuscated such that they are indistinguishable from each other and can cooperate to improve privacy protection.

The contributions of this chapter are summarized as follows,

1) We develop differential privacy in road networks and propose the RN-Indistinguishability that quantifies road network location privacy with unique road network features.

2) By using the RN-Indistinguishability, we design the CRO mechanism that employs the route distances to quantify the indistinguishability of locations on roads. We prove that the CRO mechanism satisfies the RN-Indistinguishability. The CRO mechanism can be

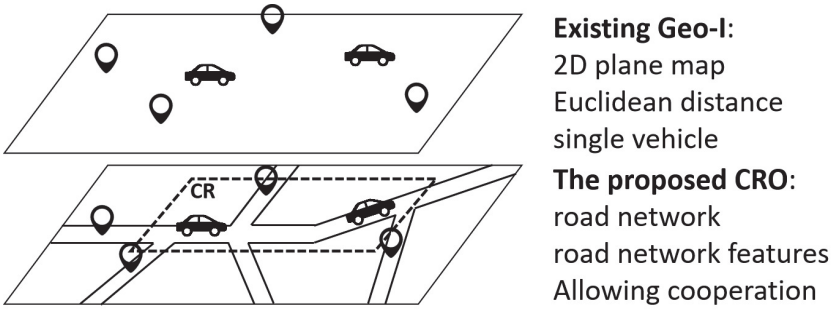


Figure 4.5: The difference of the existing Geo-I mechanism and the proposed Cloaking Region Obfuscation (CRO) mechanism.

extended with general road network features without breaching differential privacy.

We evaluate the CRO mechanism with real-world road networks in comparison with the state-of-the-art 2D Laplace location privacy-preserving mechanisms [77], [89]. Experiment results confirm the superiority of the CRO mechanism in terms of privacy-preserving level and data utility.

4.7 Related Work of Cloaking Region Obfuscation for Road Network-Indistinguishability and Location Privacy

Location privacy-preserving mechanisms can be classified into user-side mechanisms, server-side mechanisms, and channel-side mechanisms [90]. Studies mainly use obfuscation, anonymity, and cryptography to realize these three mechanisms [65]. Obfuscation mechanisms aim to reduce the precision of location data by adding noise in the actual location data [377]. Obfuscation mechanisms suit the LBSs that perform fine with coarse location data [372]. Anonymity mechanisms allow vehicles to use pseudonyms to hide vehicles' real identities [372]. However, the spatial-temporal correlation of the vehicles' trajectories can be estimated by adversaries [378]. Cryptographic algorithms can be used to encrypt the actual location data. Alternatively, certified credentials can be allocated to drivers from trusted authorities [379]. The cryptographic mechanisms have a limitation in time-sensitive LBSs due to the potentially high computational complexity and untrusted entities [380].

Differential privacy (DP) technology is widely used in obfuscation mechanisms [333]. It protects individuals' private information indistinguishable when publishing aggregated data [240], [334].

Most of the existing obfuscation mechanisms protect location privacy on two-dimensional (2D) maps. Andrés *et al.* developed Geo-Indistinguishability (Geo-I) based on differential privacy [74]. The authors employed Laplace mechanism [335] to realize Geo-Indistinguishability on a 2D plane. The Geo-I reports obfuscated locations surrounding the actual location based on the obfuscated probability distribution when a driver requests LBS. Hua *et al.* introduced additional servers to assist Geo-I obfuscation process [77]. The authors divided the map into several areas by introducing extra servers. A driver employs the actual location of the servers to calculate the obfuscated locations. Rather than dividing maps, the mechanism in [69] splits roads into same-length road sections. The authors employed the route distance between two road sections to measure the indistinguishability of the Geo-Indistinguishability. Drivers who

Table 4.2: Summary of notations and abbreviations

Notation	Description
SD	Shift Distance
AEE	Adversary Estimation Error
CRO	Cloaking Region Obfuscation
$\mathcal{G}$	The graph transformed from a road network
l	The location of a cloaking region
$d_{\mathcal{G}}(l, l')$	The shortest route distance between l and l'
$\mathcal{D}_{\mathcal{G}}(l, l')$	The indistinguishability between l and l'
ϵ	Privacy budget
$\mathbb{R}^2$	Obfuscation area
r	The radius of obfuscation areas

locate in the same road section are obfuscated in the same rule. The Graph-Exponential Mechanism (GEM) evaluates the privacy-preserving capability of the Geo-Indistinguishability in road networks [73]. The GEM sets the connections (such as turns, intersections, furcates, and joins) in a road network as the obfuscation candidates. The mechanism maps drivers' actual locations to connections and obfuscates connections based on the Geo-Indistinguishability.

The cloaking region, in which drivers' private data are uniformly processed, is introduced to spatially or temporally diminish the accuracy of the drivers' actual location data in LBS queries [222]. The existing cloaking-region mechanisms focus on anonymity and obfuscation in a 2D map rather than in road networks. Luo *et al.* improved a distributed k-anonymity (spatially) mechanism by using blockchain [17]. The authors improved the construction process of anonymous cloaking region based on the historical information recorded on public available blocks. Yang *et al.* designed a reputation calculation algorithm based on blockchain [381]. The authors used trusted nodes to create cloaking regions, in which the location data can be protected. To curb the malicious behaviors, the authors improved a punishment factor in their algorithm. Rather than blockchain, Li *et al.* protected location privacy with social intimate fogs [382]. The authors transferred a large amount of LBS requests to social intimate fogs, which improves efficiency but also increases communication delay (temporally). The authors improved a fog-based LBS agent which can re-encrypt multiple location data before transmitting to LBS. Cloaking region establishment is a reputation-computation question in the anonymity area (e.g., [17] and [381]), which is not the key point of this chapter.

To the best of our knowledge, none of the Geo-I and existing cloaking-region mechanisms obfuscate actual locations by considering the road network features, such as route distance and road network topology.

4.8 System Model of Cloaking Region Obfuscation for Road Network-Indistinguishability and Location Privacy

4.8.1 Road Network Model

Drivers' request various LBS when driving in road networks following traffic rules. Considering the untrustworthiness of the LBS servers, the drivers do not want to share their precise location data with the servers. The drivers can protect their location privacy by employing obfuscation mechanisms and reporting obfuscated locations to the servers.

Road network

A road network is modeled as a directed and weighted graph $\mathcal{G}$. The road network nodes are the connections (e.g., turns, intersections, furcates, and joints) mapped to the same geo-locations in the real world. The weight of each edge indicates the length of the corresponding road section between two connections.

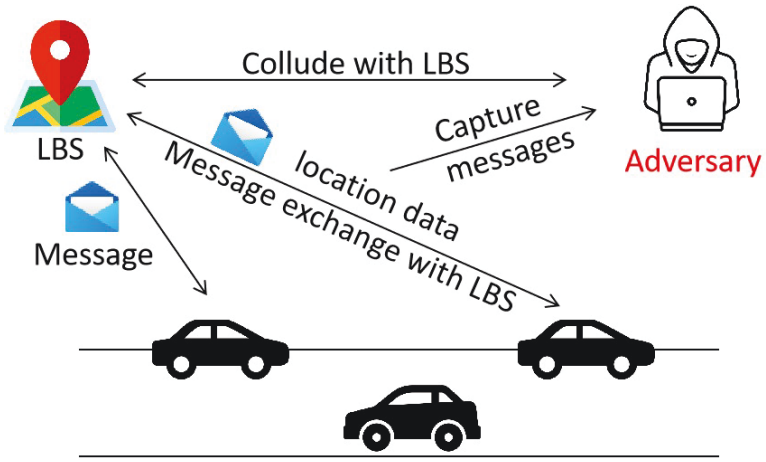


Figure 4.6: Adversary model.

Cloaking regions (CRs)

We evenly divide road networks into CRs of the same area size. A CR is empty if there is no connection or road inside it. The empty CR is removed in our model.

Actual location l

In each CR, we select the closest on-road point to the center of the CR to represent the CR. The location of the n-th CR, denoted by CR_n, is represented by the location of the selected point in CR_n, denoted by, l_n. The locations of vehicles in CR_n use l_n as their actual locations.

Obfuscation area $\mathbb{R}^2$

The obfuscation area $\mathbb{R}^2$ of l is a circular region with radius r and centered on l, in which the locations of the vehicles can be obfuscated based on the obfuscation probability. A CR_n is

inside $\mathbb{R}^2$ if its center location l_n is in $\mathbb{R}^2$. All CRs in $\mathbb{R}^2$ are obfuscation candidates of the CR with the location l.

Route distance

The route distance between a pair of CRs (CR_i and CR_j) is measured using the route distance between l_i and l_j, i.e., $d_{\mathcal{G}}(l_i, l_i)$.

LBS request

The LBS requests are sent from vehicles to the LBS servers. In our model, an unprotected LBS query q from a vehicle in a CR is as follows,

$$q = (l, I),$$

where l is the location of the CR within which the vehicle is. I is other contents of the query.

We assume that the LBS servers are not trusted and may disclose the drivers' data to adversaries. We use differential privacy to protect drivers' location privacy. When a driver's location needs to be protected, the driver can access to the LBS server by sending a private query q' with an obfuscated location, as given by

$$q' = (l', I), \tag{4.2}$$

where l' is the location of the obfuscated cloaking region selected based on l.

We define shift distance (SD) to measure the data utility in the queries affected by the obfuscation mechanism. Given a set of queries from CR with l, the average SD of the cloaking region ($SD(l)$) is given by [383, eq. 3]

$$SD(l) = \frac{1}{K} \sum_{k=1}^{K} d_{\mathcal{G}}(l, l'_k), \tag{4.3}$$

where K is the number of queries from the drivers in the CR. l'_k is the location of the obfuscated CR in the k-th query. The SD measures the data utility of obfuscated locations. A smaller SD indicates a higher data utility and a potentially better LBS. The data utility decreases when the drivers employ the obfuscation mechanisms to protect their location information.

4.8.2 Adversary Model

Both the external and internal passive attackers are considered in this chapter, as shown in Fig. 4.6. External adversaries can eavesdrop on the LBS messages exchanged between the LBS servers and the vehicles to obtain the vehicles' locations. Internal adversaries can collude with the LBS servers to obtain the vehicles' queries. We assume that the adversaries are fully aware of the drivers' prior trajectories and the obfuscation mechanism [69].

We follow the adversary model in [207]. The actual location of a driver is unknown and any other parties (e.g., adversaries and servers) cannot directly observe the driver's actual location. The drivers' identities can be protected by the existing anonymity mechanisms, so that adversaries cannot track drivers' trajectories by analyzing identity information. Other information, such as, obfuscation mechanisms and parameters, is known to all parties.

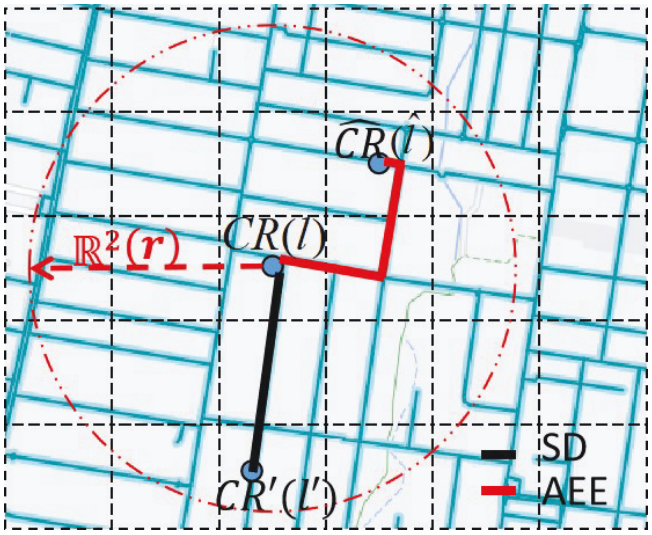


Figure 4.7: An example of the proposed CRO algorithm.

Obfuscation Probability: Given the location l of the CR within which vehicle in, location obfuscation mechanisms select an obfuscated CR with location l' based on the obfuscation probability $\Pr[l'|l]$.

The adversaries can infer the probability that the obfuscated location l' is generated by the location l in $\mathcal{G}$ is given by [207, eq. 1]

$$\Pr[l_i|l'] = \frac{\Pr[l'|l_i]\Pr[l_i]}{\sum_{l_j \in \mathbb{R}^2} \Pr[l'|l_j]\Pr[l_j]}, \tag{4.4}$$

where $\Pr[l_i]$ and $\Pr[l_j]$ are the prior probability that the driver locates at the CR_i (with location l_i) and CR_j (with location l_j) according to the prior trajectories, respectively. The prior probability is based on the driver's historical trajectory. $\Pr[l'|l_j]$ is the probability of the region CR' being selected as the obfuscated version of the actual region CR for a vehicle.

The adversaries select the CR, $\hat{CR}$ (with location $\hat{l}$), as the actual cloaking region of the driver, if $\hat{l}$ has

$$\Pr[\hat{l}|l'] = \max_{l_j \in \mathcal{G}} \Pr[l_j|l']. \tag{4.5}$$

The shortest route distance between the two cloaking regions CR (with location l) and $\hat{CR}$ (with location $\hat{l}$) measures the Adversary Estimation Error (AEE), which evaluates the protection effect of an obfuscation mechanism on location privacy. As illustrated in Fig. 4.7, the AEE is $d_{\mathcal{G}}(l, \hat{l})$ and the SD is $d_{\mathcal{G}}(l, l')$. Given a set of queries from CR with l, $\hat{CR}_k$ with $\hat{l}_k$ is the k-th derived CR, the average AEE of the CR ($AEE(l)$) is given by

$$AEE(l) = \frac{1}{K} \sum_{k=1}^{K} d_{\mathcal{G}}(l, \hat{l}_k). \tag{4.6}$$

4.9 Proposed Cloaking Region Obfuscation Mechanism

In this section, we define RN-Indistinguishability with general road network features (e.g., route distances, speed limits, and road network densities) based on the definition of Geo-I [74]. The

RN-Indistinguishability employs road network features to measure the indistinguishability of locations in road networks. By using the route distance as the metric of RN-Indistinguishability, we propose the Cloaking Region Obfuscation (CRO) mechanism to protect drivers' location privacy in road networks. The proposed CRO mechanism ensures that adversaries cannot infer the drivers' actual locations based on the obfuscated location data. We prove that the CRO mechanism satisfies the RN-Indistinguishability and can be generalized with multiple road network features.

4.9.1 Road Network-Indistinguishability

The Geo-I mechanism can protect the location privacy in 2D Euclidean spaces [74]. The Geo-I mechanism makes the actual locations Geo-indistinguishable to any adversaries by using the differential privacy to perturb the drivers' actual locations. In contrast, the RN-Indistinguishability assesses on-road privacy-preserving mechanisms, and is defined as follows.

Definition 7 (RN-Indistinguishability)**.** *A mechanism satisfies $(\epsilon, \mathbb{R}^2)$-RN-Indistinguishability if and only if, with an obfuscated location l', l' in $\mathbb{R}^2$), any location pair (l_i, l_j) in $\mathbb{R}^2$ have*

$$\frac{\Pr[l_i|l']}{\Pr[l_j|l']} \leq e^{\epsilon \mathcal{D}_{\mathcal{G}}(l_i,l_j)} \frac{\Pr[l_i]}{\Pr[l_j]}, \tag{4.7}$$

where $\mathcal{D}_{\mathcal{G}}(l_i, l_j)$ is the indistinguishability between l_i and l_j in road networks. ϵ is the privacy budget. $\Pr[l_i|l']$ and $\Pr[l_j|l']$ are the probabilities that l' is obfuscated based on l_i and l_j, respectively.

In this chapter, RN-Indistinguishability is employed to obfuscate cloaking regions. With the cloaking region CR_i and CR_j, $\mathcal{D}_{\mathcal{G}}(l_i, l_j)$ can be calculated by [384]

$$\begin{cases} \mathcal{D}_{\mathcal{G}}(l_i,l_j) = \alpha_1 \dfrac{d_{\mathcal{G}}(l_i,l_j)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}} + \alpha_2 \dfrac{v_{\mathcal{G}}(l_i,l_j)}{\max_{\mathbb{R}^2} v_{\mathcal{G}}} + \alpha_3 \dfrac{den_{\mathcal{G}}(l_i,l_j)}{\max_{\mathbb{R}^2} den_{\mathcal{G}}} + \cdots ; & (4.8) \\ \alpha_1 + \alpha_2 + \alpha_3 + \cdots = 1, & (4.9) \end{cases}$$

where the route distances $d_{\mathcal{G}}(l_i, l_j)$, the speed limit difference $v_{\mathcal{G}}(l_i, l_j)$, the road network density difference $den_{\mathcal{G}}(l_i, l_j)$, and other road network features can be considered. α_1, α_2, and α_3 are weighting factors. $\max_{\mathbb{R}^2} d_{\mathcal{G}}$, $\max_{\mathbb{R}^2} v_{\mathcal{G}}$, and
$\max_{\mathbb{R}^2} den_{\mathcal{G}}$ are the maximum values of route distance, speed limit differences, and road network density differences in $\mathbb{R}^2$, respectively.

By substituting (4.4) into (4.7), we can have

$$\begin{aligned} \frac{\frac{\Pr[l'|l_i]\Pr[l_i]}{\sum_{l_k \in \mathbb{R}^2} \Pr[l'|l_k]\Pr[l_k]}}{\frac{\Pr[l'|l_j]\Pr[l_j]}{\sum_{l_k \in \mathbb{R}^2} \Pr[l'|l_k]\Pr[l_k]}} &\leq e^{\epsilon \mathcal{D}_{\mathcal{G}}(l_i,l_j)} \frac{\Pr[l_i]}{\Pr[l_j]}; \\ \frac{\Pr[l'|l_i]\Pr[l_i]}{\Pr[l'|l_j]\Pr[l_j]} &\leq e^{\epsilon \mathcal{D}_{\mathcal{G}}(l_i,l_j)} \frac{\Pr[l_i]}{\Pr[l_j]}; \\ \frac{\Pr[l'|l_i]}{\Pr[l'|l_j]} &\leq e^{\epsilon \mathcal{D}_{\mathcal{G}}(l_i,l_j)}; \\ \Pr[l'|l_i] &\leq e^{\epsilon \mathcal{D}_{\mathcal{G}}(l_i,l_j)} \Pr[l'|l_j], \end{aligned} \tag{4.10}$$

where the two cloaking regions CR_i (with location l_i) and CR_j (with location l_j) are RN-indistinguishable to any adversaries. $\Pr[l'|l_i]$ and $\Pr[l'|l_j]$ are the probabilities of the region CR' being selected as the obfuscation version of the actual region CR_i and CR_j, respectively

4.9.2 Cloaking Region Obfuscation

In this subsection, we propose a CRO mechanism based on the RN-Indistinguishability. We set $\alpha_1 = 1$ and the other weighting factors to be 0, which indicates that we only consider $d_{\mathcal{G}}(l_i, l_j)$ as the metric of RN-Indistinguishability.

In the CRO mechanism, we obfuscate the location of the cloaking region, where the driver actually locates, and use its location in the LBS query. All the drivers in the same cloaking region are obfuscated following the same rule, and therefore are indistinguishable from each other. The cloaking regions in road networks can also prevent the adversaries from identifying and locating vehicles.

The proposed CRO mechanism consists of the following steps:

- **Initialization:** The road network is evenly divided into cloaking regions. Vehicles in the cloaking region CR (with location l) specify the obfuscation area $\mathbb{R}^2$, centering l with radius r. Cloaking region CR_i (with location l_i) is in $\mathbb{R}^2$ if l_i is inside $\mathbb{R}^2$.
- **Obfuscation:** Given a vehicle in CR, the location of the vehicle is represented by l. The vehicle probabilistically selects an obfuscated cloaking region CR' (with location l') in $\mathbb{R}^2$. The probability $\Pr[l'|l]$ is as given by

$$\Pr[l'|l] = \frac{e^{-\frac{\epsilon}{2}\frac{d_{\mathcal{G}}(l,l')}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}}}{\sum\limits_{l_i \in \mathbb{R}^2} e^{-\frac{\epsilon}{2}\frac{d_{\mathcal{G}}(l,l_i)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}}}, \tag{4.11}$$

  where l_i is the location of the i-th cloaking region in $\mathbb{R}^2$. $\max_{\mathbb{R}^2} d_{\mathcal{G}}$ is the longest route distance between two cloaking regions in $\mathbb{R}^2$. All vehicles in the same cloaking region have the same obfuscation probability.
- **Finalization:** The vehicle reports a query with l' and other information to the LBS server.

In the proposed CRO algorithm, the locations of vehicles in the same cloaking are uniformly obfuscated. Thus, the location privacy-preserving capability increases with the number of vehicles.

4.9.3 Privacy Analysis

In this subsection, we prove that the CRO algorithm with the route distance metric satisfies the RN-Indistinguishability and can protect location privacy under the adversary model in Section 4.8.2.

Theorem 2. *With route distance metric, the CRO mechanism satisfies the (ϵ, $\mathbb{R}^2$)-RN-Indistinguishability that the locations of any two cloaking regions in an obfuscation area $\mathbb{R}^2$ are (ϵ, $\mathbb{R}^2$)-RN-indistinguishable.*

Proof. The cloaking region obfuscation in the CRO mechanism is formulated as (4.11). Let

$$\begin{cases} f(l_0) = \sum\limits_{l \in \mathbb{R}^2} e^{-\frac{\epsilon}{2}\frac{d_{\mathcal{G}}(l_0,l)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}}; & (4.12) \\ f(l_1) = \sum\limits_{l \in \mathbb{R}^2} e^{-\frac{\epsilon}{2}\frac{d_{\mathcal{G}}(l_1,l)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}}. & (4.13) \end{cases}$$

Given any location pair (l_0, l_1) in $\mathbb{R}^2$, we have

$$\frac{\Pr[l'|l_0]}{\Pr[l'|l_1]} = \frac{f(l_1)}{f(l_0)} e^{\frac{\epsilon}{2}\left(\frac{d_{\mathcal{G}}(l_1,l')}{\max_{\mathbb{R}^2} d_{\mathcal{G}}} - \frac{d_{\mathcal{G}}(l_0,l')}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}\right)}, \tag{4.14}$$

where l' is the location of the obfuscated cloaking region CR'.

Then, we prove that route distance satisfies the triangle inequality in road networks, i.e., *for any three cloaking regions with locations l_a, l_b, and l_c in a road network, $d_{\mathcal{G}}(l_a, l_c) - d_{\mathcal{G}}(l_b, l_c) \leq d_{\mathcal{G}}(l_a, l_b)$ holds as follows.*

If l_b is on the shortest route between l_a and l_c, we have

$$d_{\mathcal{G}}(l_a, l_c) - d_{\mathcal{G}}(l_b, l_c) = d_{\mathcal{G}}(l_a, l_b). \tag{4.15}$$

If l_b is not on the shortest route between l_a and l_c, by assuming that $d_{\mathcal{G}}(l_a, l_c) - d_{\mathcal{G}}(l_b, l_c) > d_{\mathcal{G}}(l_a, l_b)$, we have

$$d_{\mathcal{G}}(l_a, l_c) > d_{\mathcal{G}}(l_a, l_b) + d_{\mathcal{G}}(l_b, l_c),$$

which indicates that the route between l_a and l_c is longer than the route between l_a and l_c via l_b. Then, the shortest route between l_a and l_c should pass l_b. This contradicts with the assumption of $d_{\mathcal{G}}(l_a, l_c) - d_{\mathcal{G}}(l_b, l_c) > d_{\mathcal{G}}(l_a, l_b)$. If l_b is not on the shortest route between l_0 and l_c, the following inequality holds

$$d_{\mathcal{G}}(l_a, l_c) - d_{\mathcal{G}}(l_b, l_c) \leq d_{\mathcal{G}}(l_a, l_b). \tag{4.16}$$

Combining (4.15) and (4.16), the triangle inequality in road networks is proved.

With the triangle inequality in road networks, we have

$$\begin{aligned} d_{\mathcal{G}}(l_0, l_1) = d_{\mathcal{G}}(l_1, l_0) &\geq d_{\mathcal{G}}(l_1, l') - d_{\mathcal{G}}(l_0, l'); \\ \frac{d_{\mathcal{G}}(l_0, l_1)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}} = \frac{d_{\mathcal{G}}(l_1, l_0)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}} &\geq \frac{d_{\mathcal{G}}(l_1, l')}{\max_{\mathbb{R}^2} d_{\mathcal{G}}} - \frac{d_{\mathcal{G}}(l_0, l')}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}. \end{aligned} \tag{4.17}$$

Then,

$$\frac{\Pr[l'|l_0]}{\Pr[l'|l_1]} \leq \frac{f(l_1)}{f(l_0)} e^{\frac{\epsilon}{2}\frac{d_{\mathcal{G}}(l_0,l_1)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}}. \tag{4.18}$$

By employing the triangle inequality, we further have

$$e^{-\frac{\epsilon}{2}\frac{d_{\mathcal{G}}(l_1,l')}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}} \leq e^{-\frac{\epsilon}{2}\left(\frac{d_{\mathcal{G}}(l_0,l') - d_{\mathcal{G}}(l_0,l_1)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}\right)}. \tag{4.19}$$

Therefore, for all l in $\mathbb{R}^2$, the following holds

$$\sum_{l \in \mathbb{R}^2} \left(e^{-\frac{\epsilon}{2}\frac{d_{\mathcal{G}}(l_1,l)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}} - e^{-\frac{\epsilon}{2}\left(\frac{d_{\mathcal{G}}(l_0,l) - d_{\mathcal{G}}(l_0,l_1)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}\right)} \right) \leq 0, \tag{4.20}$$

which can be rewritten as

$$\sum_{l \in \mathbb{R}^2} e^{-\frac{\epsilon}{2}\frac{d_{\mathcal{G}}(l_1,l)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}} - e^{\frac{\epsilon}{2}\frac{d_{\mathcal{G}}(l_0,l_1)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}} \sum_{l \in \mathbb{R}^2} e^{-\frac{\epsilon}{2}\frac{d_{\mathcal{G}}(l_0,l)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}} \leq 0. \tag{4.21}$$

Based on the definition of $f(l_0)$, we have

$$f(l_1) - e^{\frac{\epsilon}{2}\frac{d_{\mathcal{G}}(l_0,l_1)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}} f(l_0) \leq 0, \tag{4.22}$$

which leads to

$$\frac{f(l_1)}{f(l_0)} \leq e^{\frac{\epsilon}{2}\frac{d_{\mathcal{G}}(l_0,l_1)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}}. \tag{4.23}$$

Combining (4.18) and (4.23), we have

$$\frac{\Pr[l'|l_0]}{\Pr[l'|l_1]} \leq e^{\frac{\epsilon}{2}\frac{d_{\mathcal{G}}(l_0,l_1)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}} e^{\frac{\epsilon}{2}\frac{d_{\mathcal{G}}(l_0,l_1)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}} = e^{\epsilon\frac{d_{\mathcal{G}}(l_0,l_1)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}}, \tag{4.24}$$

which satisfies (4.10). In other words, CR_0 and CR_1 in $\mathbb{R}^2$ are $(\epsilon, \mathbb{R}^2)$-RN-indistinguishable.

□

4.9.4 Generalization with Road Network Features

We have used the route distance as the only metric of the RN-Indistinguishability in the proposed CRO mechanism. In this subsection, we prove the CRO mechanism satisfies RN-Indistinguishability, when it is generalized with other road networks features (i.e., route distances, speed limits, and road network densities).

Theorem 3. *The CRO mechanism satisfies the (ϵ,$\mathbb{R}^2$)-RN-Indistinguishability when the employed road network features satisfy the triangle inequality.*

Proof. With multiple road network features, the CRO mechanism satisfies (ϵ,$\mathbb{R}^2$)-RN-Indistinguishability if

$$\frac{\Pr[l'|l_0]}{\Pr[l'|l_1]} \leq e^{\epsilon \mathcal{D}_{\mathcal{G}}(l_0.l_1)}. \tag{4.25}$$

As shown in (4.14), the left-hand side of (4.25) can be rewritten as

$$\frac{\Pr[l'|l_0]}{\Pr[l'|l_1]} = \frac{\sum\limits_{l_j \in \mathbb{R}^2} e^{\frac{\epsilon}{2}\mathcal{D}_{\mathcal{G}}(l_j,l')}}{\sum\limits_{l_i \in \mathbb{R}^2} e^{\frac{\epsilon}{2}\mathcal{D}_{\mathcal{G}}(l_i,l')}} e^{\frac{\epsilon}{2}(\mathcal{D}_{\mathcal{G}}(l_0,l') - \mathcal{D}_{\mathcal{G}}(l_1,l'))}, \tag{4.26}$$

where we consider multiple road network features (e.g., route distances $d_{\mathcal{G}}(l,l')$, speed limits $v_{\mathcal{G}}(l,l')$, and road network density $den_{\mathcal{G}}(l,l')$) in $\mathcal{D}_{\mathcal{G}}(l,l')$, as shown in (4.8). Let

$$\mathcal{D}_{\mathcal{G}}(l,l') = \sum_{k=1}^{K} \alpha_k g_{\mathcal{G}}^k(l,l'), \tag{4.27}$$

where K is the number of considered road network features, $g_{\mathcal{G}}^k(l,l')$ is the k-th road network feature difference between l and l', and α_k is the k-th weighting factor. As the features hold the triangle inequality, we have

$$\mathcal{D}_{\mathbb{G}}(l_0,l') - \mathcal{D}_{\mathbb{G}}(l_1,l') = \sum_{k=1}^{K} \alpha_k \left(g_{\mathcal{G}}^k(l_0,l') - g_{\mathcal{G}}^k(l_1,l')\right) \leq \sum_{k=1}^{K} \alpha_k g_{\mathcal{G}}^k(l_0,l_1) = \mathcal{D}_{\mathbb{G}}(l_0,l_1). \tag{4.28}$$

By following the analysis of (4.18) – (4.23) in the proof of *Theorem* 2, the following holds

$$\begin{cases} e^{\frac{\epsilon}{2}\mathcal{D}_{\mathcal{G}}(l_0,l')-\mathcal{D}_{\mathcal{G}}(l_1,l')} \leq e^{\frac{\epsilon}{2}\mathcal{D}_{\mathcal{G}}(l_0,l_1)}; & (4.29) \\ \dfrac{\sum\limits_{l_j\in\mathbb{R}^2} e^{\frac{\epsilon}{2}\mathcal{D}_{\mathcal{G}}(l_j,l')}}{\sum\limits_{l_i\in\mathbb{R}^2} e^{\frac{\epsilon}{2}\mathcal{D}_{\mathcal{G}}(l_i,l')}} \leq e^{\frac{\epsilon}{2}\mathcal{D}_{\mathcal{G}}(l_0,l_1)}. & (4.30) \end{cases}$$

Finally, by substituting (4.29) and (4.30) into (4.26), we have

$$\frac{\Pr[l'|l_0]}{\Pr[l'|l_1]} \leq e^{\epsilon\mathcal{D}_{\mathcal{G}}(l_0,l_1)}, \tag{4.31}$$

which satisfies (4.25). In other words, CR_0 and CR_1 are RN-indistinguishable

□

Corollary 1. *The proposed CRO mechanism using route distance $d_{\mathcal{G}}(l,l')$, speed limit difference $v_{\mathcal{G}}(l,l')$, and road density difference $den_{\mathcal{G}}(l,l')$ for $\mathcal{D}_{\mathcal{G}}(l,l')$ satisfies $(\epsilon,\mathbb{R}^2)$-RN-Indistinguishability.*

Proof. The triangle inequality of route distance $d_{\mathcal{G}}(l,l')$ in road network has been proved by following the analysis of (4.18) – (4.23) in the proof of Theorem 2.

The speed limit difference $v_{\mathcal{G}}(l,l')$ and road network density difference between two cloaking regions CR (with location l) and CR' (with location l') is as follows

$$\begin{cases} v_{\mathcal{G}}(l,l') = |v(l) - v(l')|; & (4.32) \\ den_{\mathcal{G}}(l,l') = |den(l) - den(l')|, & (4.33) \end{cases}$$

where $v(l)$ and $den(l)$ are the speed limit and road network density in cloaking region CR, respectively.

For any positive number a, b, and c, the following holds

$$|a-b| - |b-c| \leq |a-c|. \tag{4.34}$$

Using the definition of $v_{\mathcal{G}}(l,l')$ in (4.32), the above equation can be rewritten as

$$\begin{gathered} |v(l_0)-v(l_2)| - |v(l_1)-v(l_2)| \leq |v(l_0)-v(l_1)|; \\ v_{\mathcal{G}}(l_0,l_2) - v_{\mathcal{G}}(l_1,l_2) \leq v_{\mathcal{G}}(l_0,l_1); \\ \frac{v_{\mathcal{G}}(l_0,l_2)}{\max_{\mathbb{R}^2} v_{\mathcal{G}}} - \frac{v_{\mathcal{G}}(l_1,l_2)}{\max_{\mathbb{R}^2} v_{\mathcal{G}}} \leq \frac{v_{\mathcal{G}}(l_0,l_1)}{\max_{\mathbb{R}^2} v_{\mathcal{G}}}, \end{gathered} \tag{4.35}$$

where $\max_{\mathbb{R}^2} v_{\mathcal{G}}$ is the maximum speed limit difference in $\mathbb{R}^2$.

Similarly, $den_{\mathcal{G}}(l,l')$ has

$$\frac{den_{\mathcal{G}}(l_0,l_2)}{\max_{\mathbb{R}^2} den_{\mathcal{G}}} - \frac{den_{\mathcal{G}}(l_1,l_2)}{\max_{\mathbb{R}^2} den_{\mathcal{G}}} \leq \frac{den_{\mathcal{G}}(l_0,l_1)}{\max_{\mathbb{R}^2} den_{\mathcal{G}}}. \tag{4.36}$$

We set values of α_1, α_2, and α_3 larger than 0, and set other factors equal to 0. Thus,

$$\begin{aligned}\mathcal{D}_{\mathcal{G}}(l_0,l_2)-\mathcal{D}_{\mathcal{G}}(l_1,l_2) &= \alpha_1\left(\frac{d_{\mathcal{G}}(l_0,l_2)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}-\frac{d_{\mathcal{G}}(l_1,l_2)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}\right)+\alpha_2\left(\frac{v_{\mathcal{G}}(l_0,l_2)}{\max_{\mathbb{R}^2} v_{\mathcal{G}}}-\frac{v_{\mathcal{G}}(l_1,l_2)}{\max_{\mathbb{R}^2} v_{\mathcal{G}}}\right)\\ &\quad+\alpha_3\left(\frac{den_{\mathcal{G}}(l_0,l_2)}{\max_{\mathbb{R}^2} den_{\mathcal{G}}}-\frac{den_{\mathcal{G}}(l_1,l_2)}{\max_{\mathbb{R}^2} den_{\mathcal{G}}}\right)\\ &\leq \alpha_1\frac{d_{\mathcal{G}}(l_0,l_1)}{\max_{\mathbb{R}^2} d_{\mathcal{G}}}+\alpha_2\frac{v_{\mathcal{G}}(l_0,l_1)}{\max_{\mathbb{R}^2} v_{\mathcal{G}}}+\alpha_3\frac{den_{\mathcal{G}}(l_0,l_1)}{\max_{\mathbb{R}^2} den_{\mathcal{G}}}\\ &\leq \mathcal{D}_{\mathcal{G}}(l_0,l_1).\end{aligned} \tag{4.37}$$

$\mathcal{D}_{\mathcal{G}}(l,l')$ with route distance $d_{\mathcal{G}}(l,l')$, speed limit difference $v_{\mathcal{G}}(l,l')$, and road density difference $den_{\mathcal{G}}(l,l')$ satisfies the triangle inequality in road network.

According to the Theorem 3, the proposed CRO mechanism using $d_{\mathcal{G}}(l,l')$, $v_{\mathcal{G}}(l,l')$, and $den_{\mathcal{G}}(l,l')$ satisfies $(\epsilon,\mathbb{R}^2)$-RN-Indistinguishability. □

4.10 Experimental Results of Cloaking Region Obfuscation for Road Network-Indistinguishability and Location Privacy

In this section, we evaluate the CRO mechanism with real-world road networks[4] and the T-Drive trajectory (10,357 divers in Beijing, China) [342]. We extract two real-world road networks with different densities from the Open Street Map[5], as shown in Fig. 4.7. We set $\alpha_1 = 1$ in (4.9), which indicates that we use route distance $d_{\mathcal{G}}(l_i,l_j)$ as the only metric of RN-Indistinguishability.

4.10.1 Location Privacy Protection in High-density Road Network

We first compare the average AEE of the CRO mechanism with the 2D Laplace mechanisms [77], [89] in a high-density road network, as shown in Fig. 4.8(a). The 2D Laplace mechanism treats a road network as a 2D plane and employs the Euclidean distance metrics. Different privacy-preserving requirements are characterized by privacy budgets of the cloaking regions, which are set to 0.1, 0.5, 1, 5, and 10 in the simulation. The obfuscation radii are set to 500 m and 600 m. To fairly illustrate the advantage of considering features of road networks, we compare our mechanism and the other two Laplace mechanisms based on the Euclidean distance and route distance. The size of considered cloaking regions is 50 m × 50 m in the high-density road network. A CR is empty if there is no connection or road in it. The empty CR is removed. Then, we obfuscate the cloaking regions with the proposed CRO mechanism and 2D Laplace mechanisms developed in [77], [89].

Experimental results show that the CRO mechanism protects location privacy better than the 2D Laplace mechanisms (i.e., [77], [89]) in road networks, as shown in Fig. 4.8. The obfuscation radii are 500 m and 600 m, respectively. The average AEE of the proposed CRO mechanism is

[4]There are 69239 connections and 80943 roads within the high-density road network (116.246898°≤ longitude ≤ 116.5089155°and 39.78669°≤ latitude ≤ 40.0447621°) with 83 connections and 96 roads per km^2.
There are 29989 connections and 33950 roads within the low-density road network (116.432589°≤ longitude ≤ 116.7437658°and 39.7189659°≤ latitude ≤ 40.025068°) with 8 connections and 9 roads per km^2.

[5]Open Street Map is an open source database of the world's geographic map. `https://www.openstreetmap.org/`

(a) The road network with high density.

(b) The road network with low density.

Figure 4.7: The experimental road networks.

less than that on the 2D Laplace mechanism. The average AEE of the mechanisms decreases as ϵ increases. With the proposed mechanism, a driver selects an obfuscated cloaking region, which is close to the actual location, with a higher probability of using the route distance as a metric than that of using Euclidean distance, due to (4.11). The obfuscation mechanism also has a high probability selecting a close cloaking region as obfuscated cloaking region in the high ϵ environment. Thus, the decreasing trend of the proposed CRO mechanism is faster than its counterparts. The average AEE of the CRO mechanism under the various ϵ values can be greater than the 2D Laplace mechanisms. The average AEE with Euclidean distance is much shorter than that with route distance. There are two reasons: Firstly, the adjacent cloaking regions in 2D maps may not be adjacent in road networks. Secondly, the empty cloaking regions in obfuscation areas are removed, so each obfuscation area covers an uneven number of cloaking regions.

The comparison of the average SD between the proposed CRO mechanism and the 2D Laplace mechanism is shown in Fig. 4.9. The radii of the obfuscation areas are 500 m and 600 m, respectively. The average SD is measured by the shortest route distance between the actual and obfuscated cloaking regions. The CRO mechanism outperforms the 2D mechanism regarding the average SD. This is due to the fact that the CRO mechanism has a high probability of selecting a close cloaking region as the obfuscated cloaking region. The average SD of the CRO mechanism fluctuates in Fig. 4.9. The reason is that the CRO mechanism are measured by the route distance in the obfuscate cloaking regions. In the same obfuscation area, two cloaking regions with a short Euclidean distance may have a considerably long route distance.

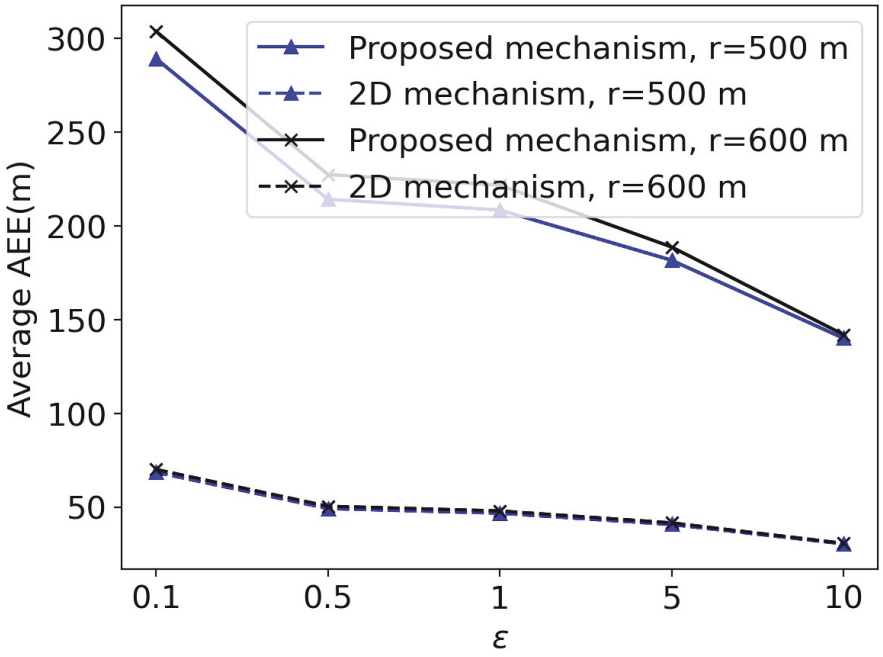
4.10.2 Location Privacy Protection in Low-density Road Network

We compare the CRO mechanism with the 2D Laplace mechanism [77], [89] in the low-density road network in Fig. 4.8(b). The size of the cloaking region is set to 50 m $\times$ 50 m. We use the 2D Laplace mechanism [77], [89] to obfuscate the cloaking regions. The privacy budgets of the cloaking regions are set to 0.1, 0.5, 1, 5, and 10 for simulating different privacy-preserving requirements. The obfuscation radii are set to 500 m and 600 m.

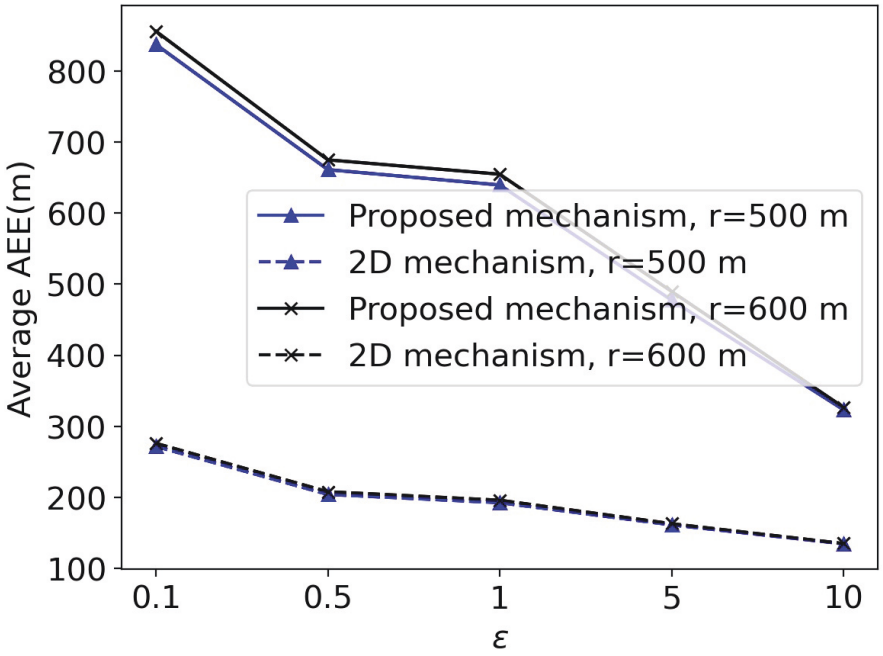
We compare the average SD of the CRO mechanism and the 2D Laplace mechanism in low-density road networks, as shown in Fig. 4.10. The CRO mechanism can achieve a shorter average SD than the 2D Laplace mechanisms in [77], [89]. The gap of the average SD with the Euclidean distance between the two curves decreases as the privacy budget increases, while that with the route distance increases with the privacy budget. The reason is that the obfuscation mechanisms have a high probability to select adjacent cloaking regions as obfuscated results. Thus, the CRO mechanism and 2D mechanisms have a similar Euclidean-distance-based average SD under a high privacy budget. The 2D mechanisms use the Euclidean distance as the metric to obfuscate cloaking regions, while the proposed CRO mechanism employs the route distance. Thus, the gap of the route-distance-based average SD between the CRO mechanism and 2D mechanisms increases with the privacy budget.

The CRO mechanism achieves a longer average AEE than the 2D mechanism in low density road networks, as shown in Fig. 4.11. The CRO mechanism and 2D mechanisms have a similar average AEE when the privacy budget is high. The average AEEs of the CRO and 2D mechanisms are shorter than the average SDs of those because we consider the worst-case scenario in the experiments, as shown in (4.6).

By comparing the results in Figs. 4.8 and 4.9 with the ones in Figs. 4.10 and 4.11, the proposed

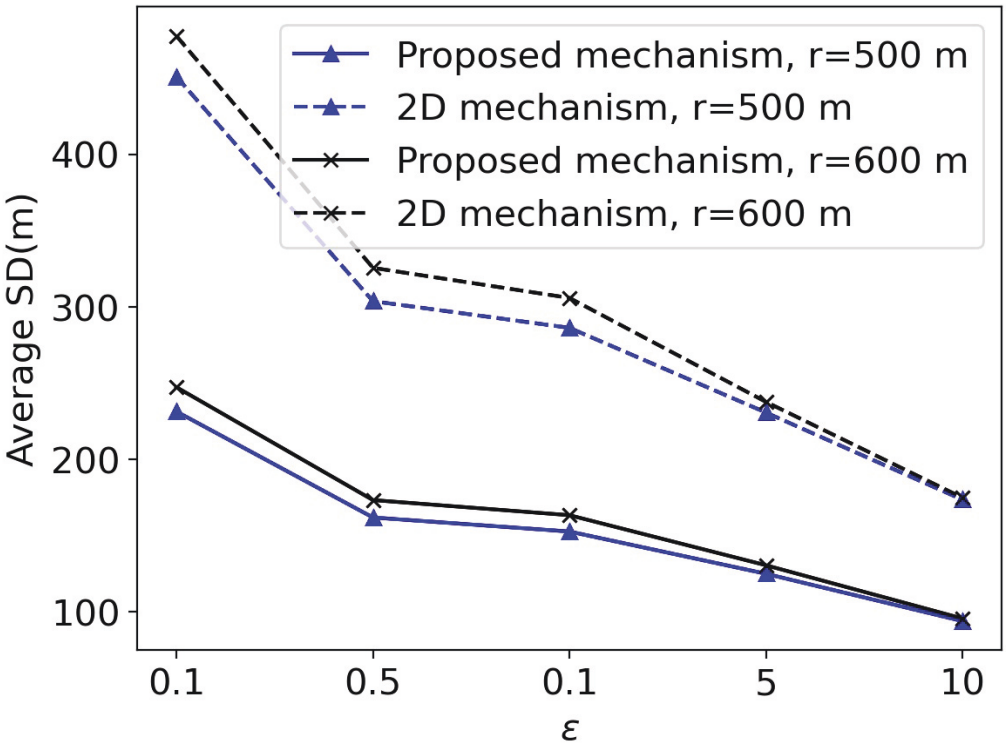


(c) Average AEE with Euclidean distance.

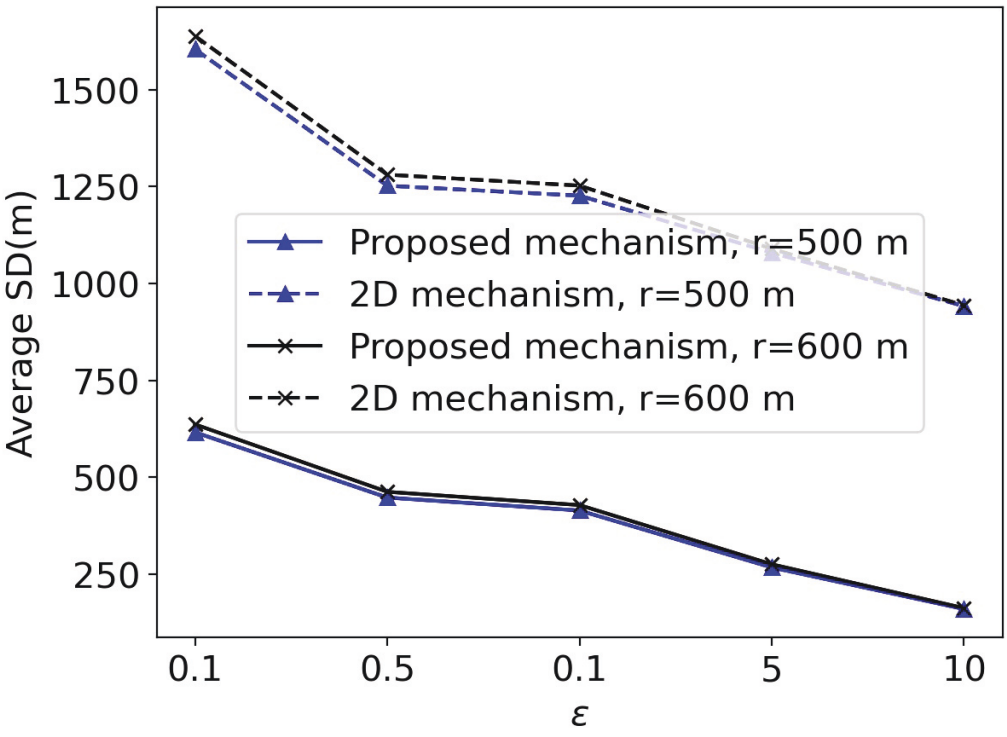


(d) Average AEE with route distance.

Figure 4.8: Average AEE comparisons of the proposed CRO mechanism and the 2D Laplace mechanisms in [77], [89] in the high-density road network.

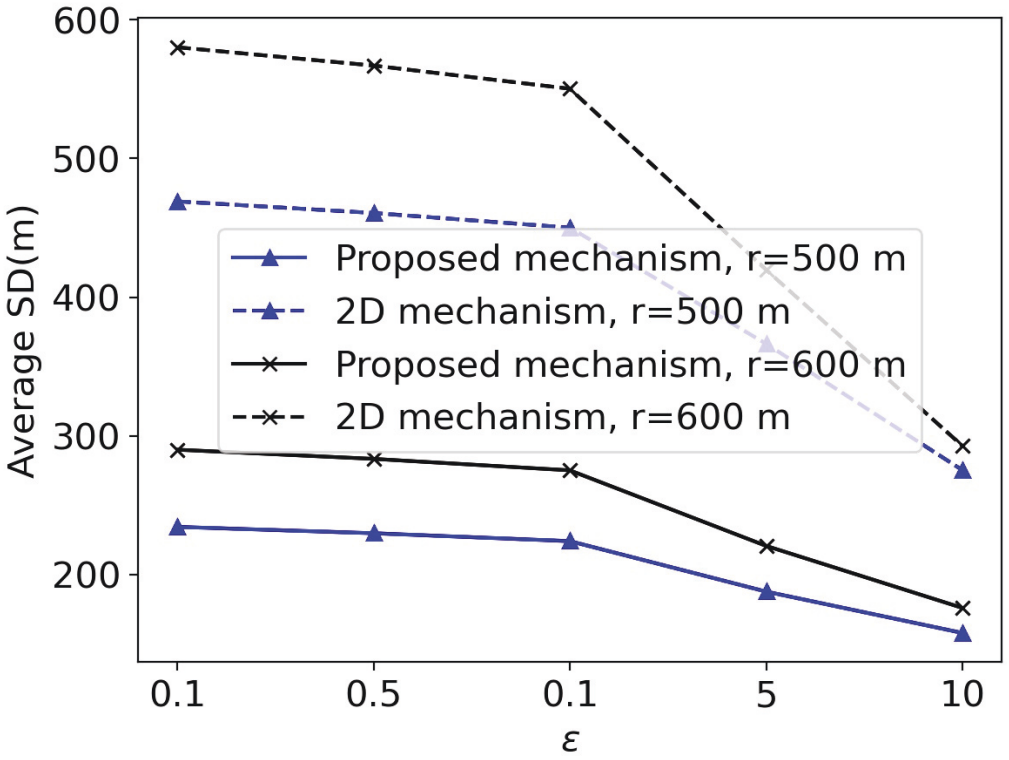


(a) Average SD with Euclidean distance.

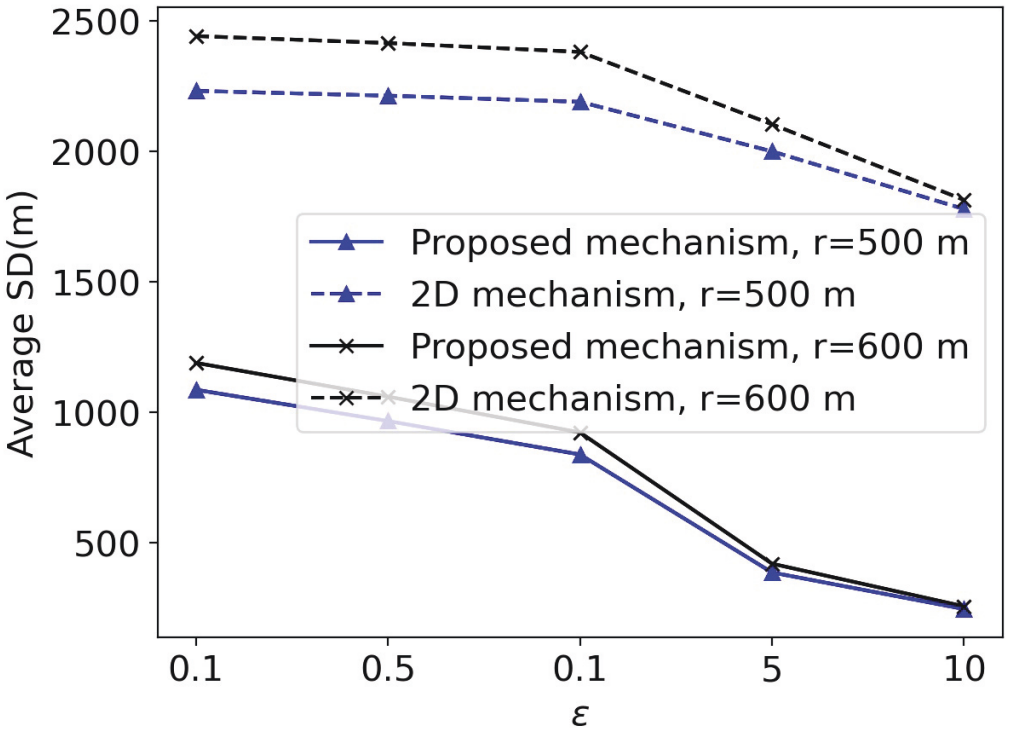


(b) Average SD with route distance

Figure 4.9: Average SD comparisons of the proposed CRO mechanism and the 2D Laplace mechanisms in [77], [89] in the high-density road network.

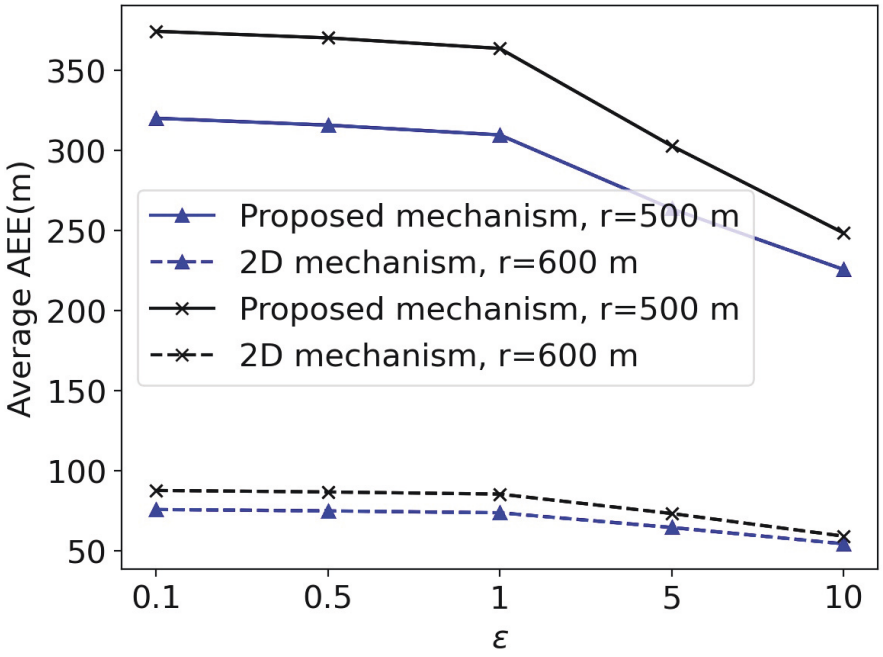


(a) Average SD with Euclidean distance

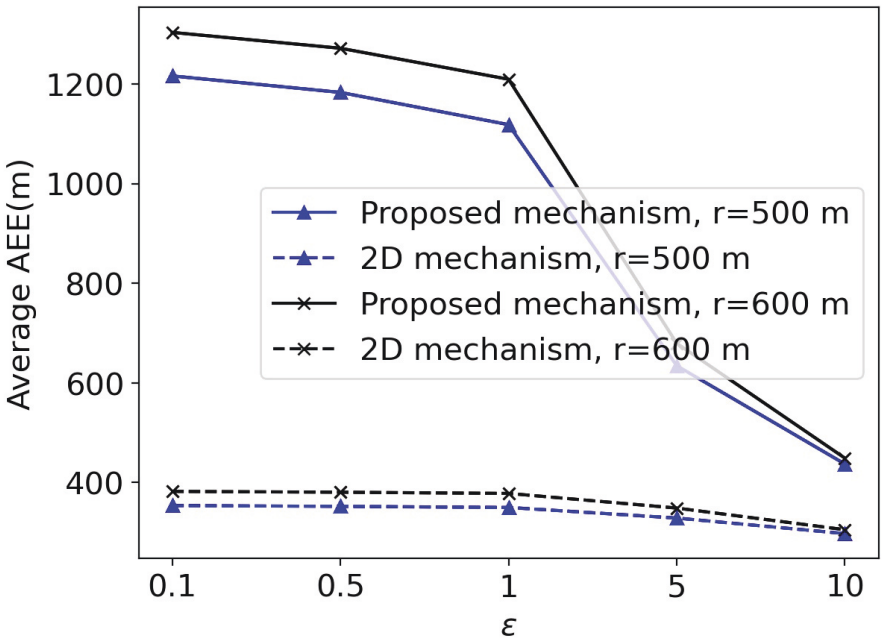


(b) Average SD with route distance

Figure 4.10: Average SD comparisons of the proposed CRO mechanism and the 2D Laplace mechanisms in [77], [89] in the low-density road network.



(a) Average AEE with Euclidean distance



(b) Average AEE with route distance

Figure 4.11: Average AEE comparisons of the proposed CRO mechanism and the 2D Laplace mechanisms in [77], [89] in the low-density road network.

CRO mechanism outperforms the 2D Laplace mechanism in the high-density and low-density road networks. The proposed CRO mechanism can better protect location privacy in high-density road networks than it does in low-density road networks. The reason is that the road network density in the high-density road networks is ten times as much as those of the low-density road networks, in our experiments. The CRO mechanism can accurately analyze the indistinguishability between cloaking regions with a complex topology and traffic conditions in road networks.

4.10.3 Generalization and Implementation

The indistinguishability of the CRO with the extended metrics, i.e., (4.8), is validated in Fig. 4.12. In this experiment, we evaluate the metrics $\mathcal{D}_{\mathcal{G}}(l_1, l_2)$ in (4.8). We set two cloaking regions l_1 and l_2 with 0 km $\leq d_{\mathcal{G}}(l_1, l_2) \leq$ 1 km, 0 km/h $\leq v_{\mathcal{G}}(l_1, l_2) \leq$ 30 km/h, and 0 connections $\leq den_{\mathcal{G}}(l_1, l_2) \leq$ 10 connections. For other parameters, we set $\max_{\mathbb{R}^2} d_{\mathcal{G}} = 1$ km, $\max_{\mathbb{R}^2} v_{\mathcal{G}} = 30$ km/h, $\max_{\mathbb{R}^2} den_{\mathcal{G}} = 10$ connections, and $\alpha_1 + \alpha_2 + \alpha_3 = 1$. And thus, we have $0 \leq \mathcal{D}_{\mathcal{G}}(l_1, l_2) \leq 1$. As shown in Fig. 4.12, the left-hand side of (4.26) is bounded that validates the indistinguishability of the CRO mechanism. An adversary can hardly distinguish the cloaking regions that have a similar $\mathcal{D}_{\mathcal{G}}(l_1, l_2)$ from the obfuscated locations, i.e., $\frac{\Pr[l'|l_1]}{\Pr[l'|l_2]} = 1$ when $\mathcal{D}_{\mathcal{G}}(l_0, l') - \mathcal{D}_{\mathcal{G}}(l_1, l') = 0$. We also see that the left-hand side of (4.26) is approximately linear with a small ϵ, e.g., $\epsilon = 0.5$. The reason is that when $0 < x << 1$, $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} \approx 1 + x$.

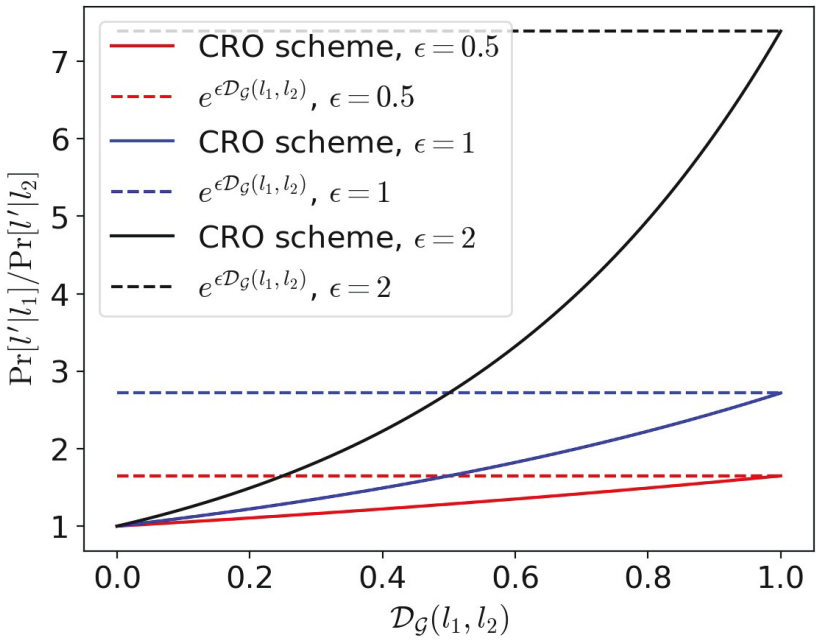


Figure 4.12: The indistinguishability of the proposed CRO mechanism with extended metrics (4.26).

The CRO mechanism is a local privacy-preserving mechanism that can independently run at every single vehicle. The CRO mechanism does not need any central server to coordinate, thereby reducing the risk of location leakage from servers or during communications. With the CRO mechanism, the locations of the vehicles in the same cloaking region are uniformly processed. The location privacy-preserving capability of the proposed CRO mechanism increases with the number of vehicles in the same cloaking region. In the case of no or few neighbors, the CRO

mechanism can still guarantee the drivers' location privacy by generating obfuscated locations because the drivers can locally run the CRO mechanism to protect their location data. The setting of the cloaking region can be optimized using the mechanisms developed in [17] and [381]. Vehicles in the same cloaking region can cooperate by swapping pseudonyms (e.g., [384]) or generating pseudonyms (e.g., [215]) to achieve high privacy protection.

4.11 Conclusion of Cloaking Region Obfuscation for Road Network-Indistinguishability and Location Privacy

In this chapter, we proposed the RN-Indistinguishability to evaluate obfuscation-based location privacy-preserving mechanisms in road networks. We proposed the new CRO mechanism to protect the location privacy of vehicles in road networks by uniformly obfuscating the locations of vehicles in the same cloaking region. The proposed CRO mechanism was proved to achieve the RN-Indistinguishability and validated with comprehensive experiments. We also proved that the CRO can be generalized with road network features without breaching the differential privacy. In the future, we will combine other location privacy-preserving mechanisms with the CRO mechanism to simultaneously perturb the vehicles' identities and locations in the road network environment.

Chapter 5

Cooperative Trajectory Privacy Protection

5.1 Introduction

Vehicular applications provide various services with locations that are uploaded from vehicles [222]. The application servers can obtain and store the geolocation and driving statuses of vehicles, raising the concern of the vehicle's location privacy [372]. Adversaries can attack the application servers and eavesdrop on communications between the servers and vehicles for shared geolocation data, with which the adversaries can infer the drivers' private information, such as whereabouts, religion, job, and home address [8], [373], [374]. The adversary can also estimate the vehicles' locations with reported driving statuses, such as driving speed and direction [385]. Compared with other mobile users, the vehicles are easier to be tracked because their trajectories are predictable in road networks [15], [375].

Obfuscation-based mechanisms have been developed to protect the location privacy of vehicles against eavesdroppers and malicious servers [72]–[74]. The vehicles can locally perturb their actual driving statuses by exploiting Differential Privacy (DP) mechanisms that add controllable noises to the data [386] or probabilistically selecting obfuscated candidates [384]. Then, the vehicles upload the obfuscated data to the servers for services [69], such as location-based recommendations [76], entertainment, crowdsensing [387], and data analysis [388]. The state-of-the-art obfuscation-based mechanisms, such as [88] and [73], have considered spatial perturbation on two-dimensional (2D) plane maps and road networks. However, the temporal information has not been well considered in road networks by the existing obfuscation mechanisms [86]. If an adversary can track the identities (ID) of the vehicles, it can reduce the perturbation noise by linking the data of a specific vehicle ID over a long period, such as a long-observation attack [87]. Hence, the existing obfuscation-based mechanisms can hardly protect location privacy on a long-term basis [88].

Pseudonym-based mechanisms allow vehicles to utilize pseudonyms, rather than actual identities, to upload messages [389], hence protecting location privacy at the time domain. A vehicle can use multiple pseudonyms to share data in different periods [390]. The pseudonym-based mechanism can be classified into pseudonym generation and pseudonym swap. The pseudonym generation requires vehicles to keep silent in a specific region and use new pseudonyms simultaneously, while the pseudonym swap asks multiple vehicles to exchange their pseudonyms. The pseudonym-based mechanisms can provide high data utility with actual data, but the limitation

is that the adversary can link multiple pseudonyms to a vehicle by using the spatial-temporal relevance among trajectories based on its prior knowledge [384], [391]. With the pseudonym swap, the vehicles exchange their pseudonyms with those which have similar trajectories, reducing the probability of accurate linkage through differential cryptanalysis compared to pseudonym generation [13]. Existing studies have shown that the more similar two vehicles' driving statuses are, the better privacy protection can be achieved [217], [392], [393].

By using existing hybrid mechanisms, vehicles report obfuscated locations generated based on their actual locations with pseudonyms [394], [395]. Pseudonym swapping relies on the spatial and temporal correlation among multiple vehicles, allowing only nearby vehicles to swap pseudonyms for location privacy and data utility. On the other hand, local obfuscation only uses the local data of a single vehicle and overlooks the spatial and temporal correlation among vehicles. It has been generally believed that pseudonym and obfuscation are two separate technologies [396], [397]; i.e., trajectory obfuscation and pseudonym swapping are processed separately and in parallel [394], [395]. Existing hybrid mechanisms separately employ pseudonym swapping and trajectory obfuscation, and typically overlook the prior knowledge in their pseudonym-swapping processes. The pseudonym-swapping process of the existing hybrid mechanisms could undergo a low data utility and privacy-preserving capability. For example, the pseudonym-swapping process without using differential privacy can be attacked by the adversary with prior knowledge [384]. As a matter of fact, we reveal analytically in this chapter that pseudonym swaps can be considered a differential privacy process for any two identities sharing the same pseudonym-swapping candidate set.

In this chapter, we propose a novel Joint Trajectory Obfuscation and Pseudonym Swapping (JTOPS) mechanism to prevent adversaries from learning actual driving statuses and matching their prior knowledge of vehicles. We start by designing a new distance metric to measure the difference among spatio-temporal driving statuses, and defining Trajectory-Indistinguishability (T-I) to evaluate the privacy-preserving ability against adversaries who have the prior knowledge of the drivers' historical trajectories. DP is leveraged to jointly obfuscate the driving statuses and swap the drivers' pseudonyms. The vehicles firstly obfuscate their driving statuses, according to the distance between their actual statuses and candidates. Then, they swap their pseudonyms, according to the swapping probabilities calculated based on the distances among their obfuscated statuses. The JTOPS is proven to achieve T-I under the Global Passive Adversaries (GPAs) model with the full prior knowledge of the vehicles at the adversary. Additionally, the proposed JTOPS combines two differential privacy processes, without introducing the additivity composition theorem of multiple ε-DP processes [398]. The proposed mechanism can still effectively disguise the vehicles' driving statuses when their swapped pseudonyms are exposed.

The contributions of this chapter are summarized as follows,

1) This chapter demonstrates that pseudonym swaps can be considered a differential privacy process for any two identities sharing the same pseudonym-swapping candidate set. We analytically prove that jointly using pseudonym swapping and obfuscation in vehicular networks can achieve higher privacy protection than separately using them.

2) We introduce a new unified privacy-preserving measure, i.e., T-I, treating trajectory obfuscation and pseudonym-based mechanisms as a holistic process. T-I extends the applicability of ε-differential privacy (DP) by quantifying the distinguishability between vehicles based on their historical information in the time domain and current information in the spatial domain.

3) Building upon T-I, we propose a novel Joint Trajectory Obfuscation and Pseudonym Swapping (JTOPS) mechanism joining pseudonym swapping and trajectory obfuscation with a crafted criterion. The proposed mechanism is proven to combine two differential privacy processes, without introducing the additivity composition theorem of multiple ε-DP processes.

4) The proposed pseudonym swapping does not require vehicles to disclose their private data, as the required parameters are computed locally and then transmitted to the coordinator. Thus, the knowledge of the coordinator, which assists in pseudonym swapping, is limited, ensuring the proposed pseudonym-swapping process can effectively resist collusion attacks.

We experimentally assess the JTOPS on the real-world road network and trajectory datasets, compared to three state-of-the-art vehicle privacy-preserving mechanisms, i.e., those developed in [75], [15], and [384]. The superiority of the JTOPS is demonstrated in terms of privacy preservation and data utility in the cases of GPAs and collusion attacks.

The rest of this chapter is organized as follows. Chapter 5.2 reviews the related works. Chapter 5.3 presents the system model. Chapter 5.4 analyzes the proposed T-I and JTOPS mechanism. The experimental evaluation of the mechanism is provided in Chapter 5.5, followed by conclusions in Chapter 5.6. Notations used in the chapter are collated in Table 5.1.

5.2 Related Work

Obfuscation and anonymity have been primarily adopted to protect location privacy [65]. Obfuscation mechanisms add DP noises to actual location data or probabilistically select obfuscation candidates to reduce the precision of location data [334], while anonymity mechanisms allow vehicles to use pseudonyms to hide their real identities in different periods [372].

Table 5.1: Summary of Notions

Notion	Description
$M, \hat{M}$	A sequence of actual/obfuscated driving statuses
$m, \hat{m}$	An actual/obfuscated driving status
$\bar{\mathcal{M}}$	The sequence of obfuscation candidate
$\bar{m}$	An obfuscation candidate
ε	Privacy budget
id	The pseudonym of vehicle
α_i	The i-th weight factor
T	The time of pseudonym swap
T-	The period of time before time T
$T+$	The period of time after time T
$(\hat{M}_a, id_a)$	The uploaded messages of v_a
$\mathcal{S}$	The pseudonym pool
β	Any driving status
λ	Data Utility
γ	Pseudonym Utility

Most existing obfuscation-based mechanisms protect actual information by utilizing DP, making

private information indistinguishable to adversaries [399]. Andrés *et al.* [74] developed Geo-Indistinguishability (Geo-I) based on DP, which employed the Laplace mechanism to realize Geo-I on a two-dimensional (2D) plane. The Geo-I reports obfuscated locations surrounding the actual location based on the obfuscation probability distribution, but the Geo-I can generate off-road obfuscated locations that compromises the data utility. The Graph-Exponential Mechanism (GEM) evaluates the privacy-preserving capability of the Geo-I in road networks [73]. The GEM sets the connections (such as turns, intersections, bifurcates, and joins) in a road network as the obfuscation candidates. The mechanism maps the vehicles' locations to the connections and obfuscates the connections based on the Geo-I. Inspired by [74] and [73], Ma *et al.* [15] improved the Geo-I to Road Network-Indistinguishability (RN-I) by using route distance between two road network locations, with which the vehicles on road networks are indistinguishable to an adversary. These existing obfuscation-based mechanisms only use location data to measure the distance between two vehicles, yet overlook other road network spatial-temporal driving statuses. Moreover, the existing obfuscation-based mechanisms allow the vehicles to use their actual identities, leading to vulnerabilities to long-observation attacks with which an adversary can achieve a high accurate estimation by observing communications of the target vehicle in a long period [87].

The pseudonym-based mechanisms allow vehicles to use unique identifications, rather than actual identities, in communications, and update the identifications periodically. Pseudonym swap and pseudonym generation are the two most popular mechanisms among pseudonym-based mechanisms to update pseudonyms. The pseudonym generation mechanisms create new pseudonyms with which multiple vehicles replace their old pseudonyms, while the pseudonym swapping mechanisms assist vehicles to exchange their pseudonyms. The pseudonym swapping mechanisms, which can provide higher data utility and less storage overhead than the pseudonym generation, have gained more attention from researchers [384]. Li *et al.* [400] pointed out that pseudonym-based mechanisms can provide a high privacy-preserving capability when a spatial-temporal context is considered. The work [384] develops a pseudonym swapping mechanism by using DP, with which the vehicles swap pseudonyms probabilistically based on the driving status similarity. Therefore, the adversary cannot gain more knowledge by observing the pseudonym swapping process or referring to historical trajectories. The existing pseudonym-based mechanisms overlook the adversary with the prior knowledge of the vehicles and the privacy-preserving capability is low in such a case.

Existing hybrid mechanisms separately use obfuscation and pseudonym. Ullah *et al.* [75] developed a multiple-level mechanism by allowing vehicles to upload multiple messages with different locations and pseudonyms. The mechanism uses the actual locations of other vehicles as dummy locations. A vehicle randomly reports multiple dummy locations and the corresponding pseudonyms with its actual location and pseudonym to the LBS, indicating that the uploaded messages of the vehicle include the information of other vehicles. In the mechanism, the pseudonym and dummy location of a vehicle are selected at the same time. The work [401] separately uses pseudonym and obfuscation in a cloud-enabled internet where vehicles synchronously generate new pseudonyms and then report the actual or obfuscated locations. The authors of [401] only discuss the pseudonym generation algorithm which is independent of the obfuscation.

In vehicular networks, cryptography-based mechanisms are commonly used due to their lightweight encryption that minimizes delays and mitigates negative impacts such as frequent disconnections, instability, and handshake failures caused by the unique features of vehicles [402], [403]. To reduce the communication consumption of zero-knowledge proof (ZKP), Li et al. proposed

an improved interactive ZKP using blockchain network management [404]. The authors presented a blockchain gateway that enables switching of vehicles between adjacent blockchains without revealing sensitive data, thereby enhancing privacy protection. Khodaei et al. developed a mix-zone-based mechanism that encrypts vehicle information within a zone using a set of virtual vehicles [405]. This mechanism enhances privacy protection with limited additional computation and communication consumption. However, existing cryptography-based mechanisms may not be suitable for high-mobility real-time Location-Based Services (LBS) in vehicular networks, as low latency and delay in vehicle-to-everything (V2X) communication are crucial requirements [406].

To the best of our knowledge, our proposed mechanism is the first joint mechanism that protects location privacy with both pseudonyms swapping and trajectory obfuscation in road networks. In the JTOPS, the pseudonyms of vehicles are swapped following the probabilities generated based on the obfuscated driving statuses. By jointly using trajectory obfuscation and pseudonym swapping, the JTOPS provides excellent location privacy protection under GPAs with the prior knowledge of vehicles or under collusion attacks.

5.3 System Model

5.3.1 Vehicular Network

A vehicular network consists of three types of entities: Trusted Authority (TA), coordinators, and vehicles, as shown in Fig. 5.1.

Trusted Authority (TA)

We assume that the TA, which has unlimited resources, is managed by state representatives [407]. The TA generates a legal pseudonym for each vehicle and stores its vehicle-pseudonym mapping for potential future accident forensics and investigation.

Coordinators

The coordinators act as intermediate communication interfaces between the vehicles and the TA [407]. In our model, the coordinators assist in the pseudonym swapping process within their network coverage regions, and upload the new pseudonym-vehicle mapping to TA. The coordinators, which might collude with adversaries, are not trusted.

- **Coverage area $\mathbb{R}$:** Each coordinator covers a region $\mathbb{R}$ in road networks, e.g., the circle in Fig. 5.1, which does not overlap with other coverage regions. The vehicles in $\mathbb{R}$ swap their pseudonyms and share the same obfuscation candidates to protect their location privacy.
- **Pseudonym pool $\mathcal{S}$:** When a coordinator assists the vehicles in swapping pseudonyms, the coordinator puts the pseudonyms of the vehicles that participate in the pseudonym swapping into a pseudonym pool $\mathcal{S}$. The number of the vehicles in $\mathcal{S}$ is no larger than that of vehicles in $\mathbb{R}$. Before the pseudonym swapping, the coordinator records the recently uploaded driving statuses of the vehicles in $\mathcal{S}$ to ensure the recorded driving statuses of each vehicle have the same length.

Vehicles

The vehicles wish to enjoy location based services, yet distrust the service providers. The vehicles employ pseudonym swapping and trajectory obfuscation to protect their identities and trajectory information, respectively. The vehicles share their obfuscated driving statuses (e.g., locations, speeds, and directions) with the coordinators.

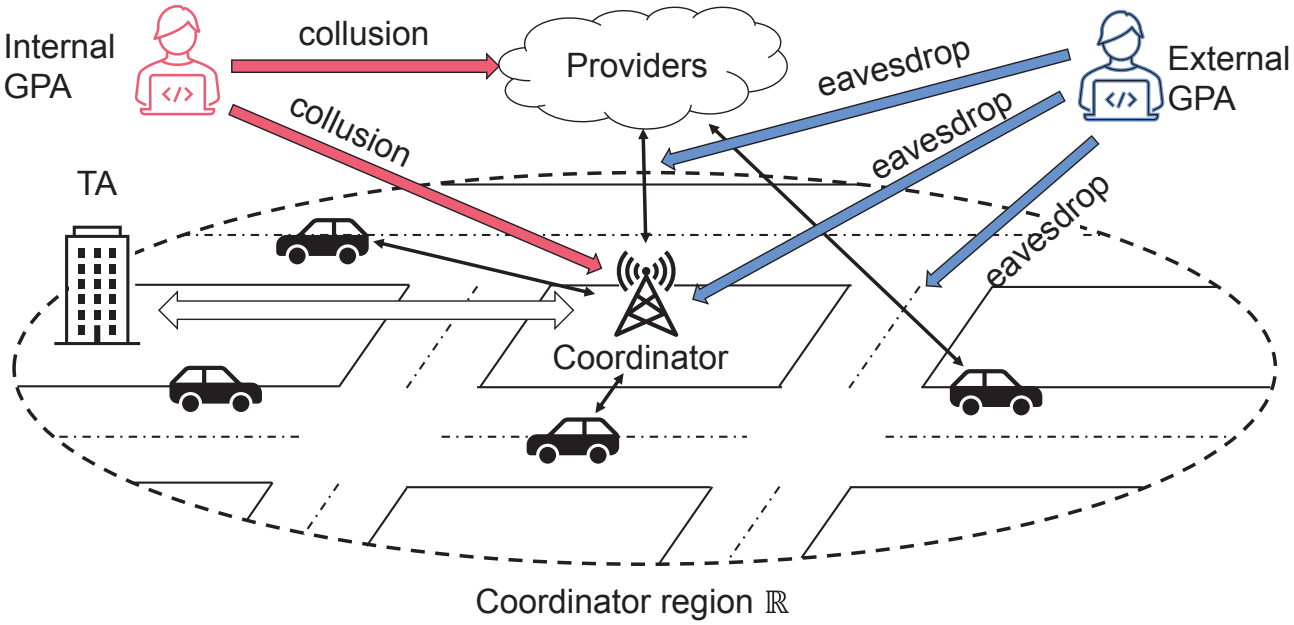


Figure 5.1: System and adversary model. The vehicular network consists of three parties: TA, coordinator, and vehicles.

A vehicle v uploads messages that contain time-series spatio-temporal driving statuses, for driving security, driving assistant, and entertainment applications [407]. The spatio-temporal information is protected by the trajectory obfuscation mechanism and uploaded with the pseudonyms. The information in the uploaded messages M is as follows:

- **Spatio-temporal driving statuses:** In this chapter, we consider the spatio-temporal driving statuses of the vehicles, e.g., location, speed, direction, and the driving time of v in $\mathbb{R}$. The uploaded messages contain multiple driving statuses at different times and are denoted by $M = \{m_1, \cdots, m_i, \cdots, m_k\}$, where $m_i \in M$ is the i-th driving statuses at time t and contains status like location loc, speed spe, and direction dir, i.e., $m_i = \{loc,\ spe,\ dir, \cdots\}$.
- **Pseudonym:** The unique identification for verification.

The uploaded messages are denoted by (M, id), where M is spatio-temporal driving statuses of vehicle v and id is the identification that v uses to upload its messages.

Relative distance between two statuses $d(M_a, M_b)$

The relative distance of driving statuses M_a (from v_a) and M_b (from v_b), denoted by $d(M_a, M_b)$, is calculated as

$$d(M_a, M_b) = \sum_i \alpha_i \frac{\beta_i(M_a, M_b)}{\max_{\mathbb{R}} \beta_i}, \tag{5.1}$$

where $\beta_i(M_a, M_b)$ denotes the distance of the i-th considered the spatio-temporal driving statuses between the two uploaded messages. Here, α_i is the i-th weighting factor, which can be set, e.g., by expert knowledge of the system designer or by machine learning[1], as long as $\sum_{\forall i} \alpha_i = 1$. $\max_{\mathbb{R}} \beta_i$ represents the maximum distance of the i-th driving statuses of the vehicles in $\mathbb{R}$. The calculation of each considered driving status is as follows.

Corollary 2. *$\mathcal{D}(v_a, v_b)$ with location distance $\beta_l(M_a, M_b)$, speed distance $\beta_s(M_a, M_b)$, direction distance $\beta_r(M_a, M_b)$, and driving time distance $\beta_t(M_a, M_b)$ satisfies triangle inequality in a road network.*

Proof. First, we prove that the vector-type driving statuses, such as route distance, satisfy the triangle inequality in road networks, i.e., *for any three vehicles v_a, v_b, and v_c in a road network, $\beta_l(M_a, M_c) - \beta_l(M_b, M_c) \leq \beta_l(M_a, M_b)$ holds as follows.*

If v_b is along the shortest route between v_a and v_c, we have

$$\beta_l(M_a, M_c) - \beta_l(M_b, M_c) = \beta_l(M_a, M_b). \tag{5.2}$$

If v_b is not along the shortest route between v_a and v_c, by assuming $\beta_l(M_a, M_c) - \beta_l(M_b, M_c) > \beta_l(M_a, M_b)$, we have

$$\beta_l(M_a, M_c) > \beta_l(M_a, M_b) + \beta_l(M_b, M_c), \tag{5.3}$$

which indicates that the route between v_a and v_c is longer than the route between v_a and v_c via v_b. Then, the shortest route between v_a and v_c should pass v_b. This contradicts with the assumption of $\beta_l(M_a, M_c) - \beta_l(M_b, M_c) > \beta_l(M_a, M_b)$. If v_b is not along the shortest route between v_a and v_c, the following inequality holds

$$\beta_l(M_a, M_c) - \beta_l(M_b, M_c) \leq \beta_l(M_a, M_b). \tag{5.4}$$

Combining (5.2) and (5.4), the triangle inequality of route distance in road networks is proved, i.e.,

$$\frac{\beta_l(M_a, M_c)}{\max_{\mathbb{R}} \beta_l} - \frac{\beta_l(M_b, M_c)}{\max_{\mathbb{R}} \beta_l} \leq \frac{\beta_l(M_a, M_b)}{\max_{\mathbb{R}} \beta_l}, \tag{5.5}$$

where $\max_{\mathbb{R}} \beta_l$ is the maximum route distance in region $\mathbb{R}$.

The numeric type of driving statuses, such as speed distance $\beta_s(M_a, M_b)$, direction distance $\beta_r(M_a, M_b)$, and the driving time distance $\beta_t(M_a, M_b)$ between any two vehicles v_a and v_b is considered as follows

$$\beta_j(M_a, M_b) = |\beta_j(M_a) - \beta_j(M_b)|, \tag{5.6}$$

where $\beta_j(M_a)$ is the j-th numeric-type driving status of vehicle v_a.

Based on the absolute value inequality and (5.6), it follows

$$\begin{gathered} |\beta_j(M_a) - \beta_j(M_b)| - |\beta_j(M_b) - \beta_j(M_c)| \leq |\beta_j(M_a) - \beta_j(M_b)|; \\ \beta_j(M_a, M_b) - \beta_j(M_b, M_c) \leq \beta_j(M_a, M_b), \end{gathered} \tag{5.7}$$

where M_c is a sequence of driving statuses of any vehicle v_c.

[1] In the experiments, all alpha values are set to be equal.

Thus, we have

$$\frac{\beta_j(M_a, M_c)}{\max_{\mathbb{R}} \beta_j} - \frac{\beta_j(M_b, M_c)}{\max_{\mathbb{R}} \beta_j} \leq \frac{\beta_j(M_a, M_b)}{\max_{\mathbb{R}} \beta_j}, \tag{5.8}$$

where $\max_{\mathbb{R}} \beta_j$ is the maximum distance of β_j in $\mathbb{R}$. Thus,

$$\mathcal{D}(v_a, v_c) - \mathcal{D}(v_b, v_c) \leq \sum_i \alpha_i \frac{\beta_i(M_a, M_b)}{\max_{\mathbb{R}} \beta_i} = \mathcal{D}(v_a, v_b). \tag{5.9}$$

In other words, $\mathcal{D}(v_a, v_b)$ satisfies triangle inequality in road networks.

For any reported driving statuses $(\hat{M}_x^{T-}, \hat{M}_x^{T+})$ from any vehicle v_c, as the driving statuses hold the triangle inequality and $0 \leq \mathcal{D}(v_a, v_b) \leq 1$, we have

$$\begin{cases} \mathcal{D}(v_b, v_c) - \mathcal{D}(v_a, v_c) \leq \mathcal{D}(v_a, v_b); \\ \mathcal{D}(v_a, v_c) - \mathcal{D}(v_b, v_c) \leq \mathcal{D}(v_a, v_b). \end{cases} \tag{5.10}$$

□

In this chapter, we consider the relative distance between the spatio-temporal driving statuses, i.e., the location distance $\beta_l(M_a, M_b)$, the speed distance $\beta_s(M_a, M_b)$, the direction distance $\beta_r(M_a, M_b)$, and the driving time distance $\beta_t(M_a, M_b)$.

5.3.2 Trajectory Obfuscation and Pseudonym Swapping

We jointly utilize pseudonym swapping and trajectory obfuscation to protect the identity and location privacy of vehicles, where the vehicles swap their pseudonyms according to the trajectory obfuscation results. When the vehicles drive in $\mathbb{R}$, they locally obfuscate their actual driving statuses for location privacy. If a vehicle needs to swap its pseudonym, it uploads its current pseudonym to $\mathcal{S}$. The vehicle then calculates the distance between its actual driving statuses and those of the other vehicles. Given the distance, the coordinator assists the vehicles in swapping their pseudonyms and returns the swapped pseudonyms to the vehicles. The vehicles use the new pseudonyms to upload their obfuscated driving statuses until the next pseudonym swap.

5.3.3 Adversary Model

We adopt the adversary model from previous studies [15], [384], and [207], which includes both external and internal GPAs as depicted in Fig. 1. These GPAs can infer the driving statuses of the vehicles by analyzing the obtained messages, and subsequently tracking the vehicles. External GPAs can operate independently of the vehicular networks and eavesdrop on the uploaded messages, whereas internal GPAs are part of the vehicular networks and may collude with the servers to obtain the uploaded messages. The GPAs have the complete knowledge of the historical driving statuses of the vehicles, the trajectory obfuscation strategy, and the pseudonym-swapping strategy.

To accurately track a vehicle, an adversary can infer the actual driving statuses of the vehicle and then estimate the real identity of the vehicle. The obfuscated driving statuses of each pseudonym can be obtained by the adversary. By linking the new and old pseudonyms, the adversary can have a set of obfuscated driving statuses of a target, i.e., $\hat{M}$. With $\hat{M}$, the

adversaries can infer the probability that $\hat{M}$ is obfuscated by the actual driving statuses M in $\mathbb{R}$. The adversary then can match the target vehicle with its prior knowledge and the inferred probability.

By considering the adversary model and the message processing, we use the following three metrics to evaluate the proposed mechanism.

Adversary's Success Rate

The Adversary's Success Rate (ASR), i.e., the probability of successfully linking the uploaded messages $\left((\hat{M}_x^{T+}, id_x^{T+}), (\hat{M}_x^{T-}, id_x^{T-})\right)$ to v_a with pseudonym id_a, is given by

$$\begin{aligned} &\Pr[id_a|((\hat{M}_x^{T+}, id_x^{T+}), (\hat{M}_x^{T-}, id_x^{T-}))] \\ =&\frac{\Pr[((\hat{M}_x^{T+}, id_x^{T+}), (\hat{M}_x^{T-}, id_x^{T-}))|id_a]}{\sum\limits_{id_c \in \mathcal{S}} \Pr[((\hat{M}_x^{T+}, id_x^{T+}), (\hat{M}_x^{T-}, id_x^{T-}))|id_c]}, \end{aligned} \tag{5.11}$$

where $(\hat{M}_x^{T-}, id_x^{T-})$ and $(\hat{M}_x^{T+}, id_x^{T+})$ are any messages in $\mathbb{R}$ before and after swapping the pseudonyms at time T. id_c is the pseudonym of any vehicle $v_c \in \mathcal{S}$.

Data Utility

We utilize the driving statuses to measure the data utility of the obfuscated driving statuses. Given a set of the obfuscated driving statuses $\hat{M}_a = \{\hat{m}_1, \cdots, \hat{m}_k\}$ from a vehicle v_a with the actual driving statuses $M_a = \{m_1, \cdots, m_k\}$, the Average Data Utility (ADU) of vehicle v_a is

$$\lambda(v_a) = \frac{1}{k}\left(1 - d(M_a, \hat{M}_a)\right), \tag{5.12}$$

where $0 \leq \lambda(v_a) \leq 1$. A high $\lambda(v_a)$ indicates that the obfuscated statuses are closer to the actual statuses; in other words, $\hat{M}_a$ provides a high data utility and quality of services.

Pseudonym Utility

The Pseudonym Utility (PU) has been overlooked in the existing pseudonym-based mechanisms. For pseudonym generation, the PU of a pseudonym is zero because the old pseudonym is abandoned. For pseudonym swap, the PU of a pseudonym could be measured by the distance of driving statuses after swapping between the two vehicles which use the pseudonym before and after swapping. A high PU indicates that the data uploaded by the pseudonym is similar, which means a high privacy-preserving capability and data utility of driving statuses. The PU can be defined as follows.

Given a pseudonym id_x, v_a is the vehicle which uses id_x before the pseudonym swapping and v_b is the vehicle that uses id_x after the pseudonym swapping. The PU of id_x is given by

$$\gamma(id_x) = 1 - d(\hat{M}_a^{T+}, \hat{M}_b^{T+}), \tag{5.13}$$

where $\hat{M}_a^{T+}$ and $\hat{M}_b^{T+}$ are the obfuscated driving statuses of v_a and v_b after swapping the pseudonyms at time T.

In (5.13), the pseudonym utility is calculated based on the driving statuses after the pseudonym swapping of the two vehicles by considering the time-sequence trajectory. For example, if a

vehicle does not change its pseudonym (id_x), the pseudonym utility would be 1, as the trajectory with the pseudonym id_x is reported as expected. However, if the pseudonym id_x is swapped from vehicle v_a to vehicle v_b, the pseudonym utility should be calculated based on the similarity between the new trajectory information of v_b (after the pseudonym swapping, i.e., $T+$) and what is expected based on the new trajectory information of v_a (after the pseudonym swapping, i.e., $T+$). The coordinator can flexibly decide the period of the pseudonym swapping using various methods, e.g., machine learning, to find the optimum value of T.

5.4 Joint Trajectory Obfuscation and Pseudonym Swapping Mechanism

In this chapter, the proposed driving status obfuscation mechanism is based on differential privacy, which has a privacy threshold due to the additivity composition theorem of reusing ε-DP. If a vehicle obfuscates its driving statuses multiple times, the privacy threshold can be exhausted. The coordinator periodically broadcasts pseudonym-swapping requests, and vehicles can decide whether to join the pseudonym swap based on their requirements. If a vehicle's privacy threshold is nearly exhausted, it will join the pseudonym-swapping process upon receiving the broadcast from the coordinator. The coordinator retains all pseudonyms until the vehicle leaves the region, while the vehicle only needs to keep its current pseudonym until it joins the next pseudonym-swapping process. Once the pseudonym of a vehicle is swapped, it keeps its new pseudonym, and the coordinator uploads the new pseudonym-vehicle mapping information of the vehicle to the Trusted Authority (TA). This process ensures that the pseudonym-vehicle mapping information is updated in a timely manner and the privacy of the vehicles is maintained throughout the pseudonym-swapping process.

5.4.1 Trajectory-Indistinguishability

The T-I considers the general spatio-temporal driving statuses of continuous trajectories in vehicular networks. The proposed T-I generalizes the metric of ε-DP in vehicular networks without increasing the upper bound of the distinguishability between two vehicles and is defined as follows.

Definition 8 (T-I). *A mechanism satisfies $(\varepsilon, \mathbb{R})$-T-I if and only if, for any two vehicles v_a and v_b in $\mathbb{R}$, which swap their pseudonyms with other vehicles in $\mathbb{R}$, the probabilities that the uploaded message $(\hat{M}_x^{T+}, id_x^{T+})$ generated by v_a and v_b have*

$$\frac{\frac{\Pr[(M_a^{T-}, id_a^{T-})|(\hat{M}_x^{T+}, id_x^{T+})]}{\sum_{v_c \in \mathbb{R}} \Pr[(M_c^{T-}, id_c^{T-})|(\hat{M}_x^{T+}, id_x^{T+})]}}{\frac{\Pr[(M_b^{T-}, id_b^{T-})|(\hat{M}_x^{T+}, id_x^{T+})]}{\sum_{v_c \in \mathbb{R}} \Pr[(M_c^{T-}, id_c^{T-})|(\hat{M}_x^{T+}, id_x^{T+})]}} \leq e^{\varepsilon \mathcal{D}(v_a, v_b)} \frac{\Pr[\![M_a^{T-}, id_a^{T-}]\!]}{\Pr[\![M_b^{T-}, id_b^{T-}]\!]}, \tag{5.14}$$

where ε is the privacy budget, and pseudonym swapping happens at time T. M_a^{T-} and M_b^{T-} are the actual driving statuses of v_a (with pseudonym id_a^{T-}) and v_b (with pseudonym id_b^{T-}) before T, respectively. id_x is any pseudonym in $\mathbb{R}$, and $\hat{M}_x^{T+}$ are the driving statuses uploaded with id_x. $\Pr[(M_a^{T-}, id_a^{T-})]$ is the probability that the actual driving statuses generated by v_a have appeared in the historical information according to the prior knowledge [408]. The prior knowledge could be gained by observing the vehicles before they are protected. (M_c^{T-}, id_c^{T-}) is the actual message

before a pseudonym swap generated by any vehicle $v_c \in \mathbb{R}$. $\mathcal{D}(v_a, v_b)$ *specifies the distance between* v_a *and* v_b*, as given by*

$$\mathcal{D}(v_a, v_b) = d(M_a^{T+}, M_b^{T+}) + d(M_a^{T-}, M_b^{T-}). \tag{5.15}$$

Here, $d(M_a^{T-}, M_b^{T-})$ *and* $d(M_a^{T+}, M_b^{T+})$ *are the driving status distance of the two vehicles before and after* T*, respectively.*

Based on Bayes' theorem, (5.14) can be written as

$$\frac{\Pr[(\hat{M}_x^{T+}, id_x^{T+})|(M_a^{T-}, id_a^{T-})]}{\Pr[(\hat{M}_x^{T+}, id_x^{T+})|(M_b^{T-}, id_b^{T-})]} \leq e^{\varepsilon \mathcal{D}(v_a, v_b)}. \tag{5.16}$$

In other words, for any sequence of messages $(\hat{M}_x^{T+}, id_x^{T+})$ after a pseudonym swap, v_a and v_b are indistinguishable.

5.4.2 Mechanism Details

In JTOPS, the vehicles locally obfuscate the spatio-temporal driving statuses. Periodically, the vehicles swap their pseudonyms with the assistance of the coordinator and then use the new pseudonyms to upload their obfuscated driving statuses. The JTOPS is described as follows.

Registration

When the vehicles first join vehicular networks, they register with the TA, and the TA allocates an initial pseudonym to each vehicle through secure channels and stores the pseudonym-vehicle mappings. Once a vehicle enters a new $\mathbb{R}$, it sends its pseudonym to the coordinator for notification. The coordinator records the pseudonym and keeps it until the vehicle leaves $\mathbb{R}$.

Trajectory obfuscation

When driving in $\mathbb{R}$, the vehicles periodically upload their messages with the pseudonyms and the obfuscated driving statuses. There are two steps in the trajectory obfuscation, i.e., generate the obfuscation candidates and obfuscate the driving statuses. The trajectory obfuscation process is described as follows.

Generate Obfuscation Candidates: The set of obfuscation candidates $\bar{\mathcal{M}} = \{\bar{m}_1, \cdots, \bar{m}_n\}$ collects all possible driving statuses of roads in the region $\mathbb{R}$, and can be calculated by the coordinator using the prior knowledge of the road network. The coordinator sends the obfuscation candidates to the vehicles once the vehicles drive into the region $\mathbb{R}$.

Obfuscate Driving Statuses: Vehicle v with the actual driving statuses $M = \{m_1, \cdots, m_i, \cdots\}$ probabilistically selects an obfuscation candidate $\bar{m}_j \in \bar{\mathcal{M}}$ as its obfuscated driving statuses $\hat{m}_i$, as given by

$$\Pr[\hat{m}_i = \bar{m}_j | m_i] = \frac{e^{-\frac{\varepsilon}{2} d(m_i, \bar{m}_j)}}{\sum\limits_{\bar{m}_x \in \bar{\mathcal{M}}} e^{-\frac{\varepsilon}{2} d(m_i, \bar{m}_x)}}, \tag{5.17}$$

where $\bar{m}_x$ is an obfuscation candidate in $\bar{\mathcal{M}}$. The obfuscation of each $m \in M$ is independent. Thus, the probability that a sequence of obfuscated driving statuses $\hat{M}$ selected based on M is

given by

$$\Pr[\hat{M}|M] = \frac{e^{-\frac{\varepsilon}{2}d(M,\hat{M})}}{\sum\limits_{\hat{M}_x} e^{-\frac{\varepsilon}{2}d(M,\hat{M}_x)}}, \tag{5.18}$$

where $\hat{M}_x$ is any sequence of possible obfuscated driving statuses that consists of multiple $\bar{m} \in \bar{\mathcal{M}}$. The lengths of different $\hat{M}_x$ are the same.

Pseudonym swapping

There are three steps in the pseudonym swapping; i.e., the coordinator records the driving statuses, the vehicles calculate the distance between the pseudonyms, and the coordinator swaps the pseudonyms. The details are described as follows.

Record Driving Status: When the vehicles upload their driving statuses, the coordinator records the obfuscated driving statuses from each pseudonym before swapping the pseudonyms. Periodically, the coordinator records the recently uploaded driving statuses from each pseudonym that are of the same length. Then, the coordinator sends all of the recorded driving statuses and the swapping requests to the vehicles. The recorded driving statuses are deleted once the pseudonym swapping process is completed.

Calculate Distances between Pseudonyms: One of the pseudonym swaps takes place at time T. If vehicle v_a using the pseudonym $id_a^{T\text{-}}$ needs to swap the pseudonyms, the vehicle calculates the distance between the pseudonym $id_a^{T\text{-}}$ and any pseudonym $id_x^{T\text{-}}$, denoted by $\mathcal{V}_{x,a}^{T\text{-}}$, by comparing the reported driving statuses, as given by

$$\mathcal{V}_{x,a}^{T\text{-}} = \frac{1}{2}d(\hat{M}_x^{T\text{-}}, M_a^{T\text{-}}) - \frac{1}{2}d(\hat{M}_a^{T\text{-}}, M_a^{T\text{-}}), \tag{5.8}$$

where $\hat{M}_x^{T\text{-}}$ and $id_x^{T\text{-}}$ are the obfuscated driving statuses and the pseudonym reported by any vehicle v_x before the pseudonym swap, respectively. We employ the scale factor $\frac{1}{2}$ in (5.8) to normalize $\mathcal{V}_{x,a}^{T\text{-}}$, i.e., $-0.5 \leq \mathcal{V}_{x,a}^{T\text{-}} \leq 0.5$. The calculation of $\mathcal{V}_{x,a}^{T\text{-}}$ uses the actual driving statuses of v_a and the obfuscated driving statuses of other vehicles, since v_a only knows the actual driving statuses of itself.

Swap Pseudonyms: Vehicles in $\mathcal{S}$ randomly select their new pseudonyms from the pseudonym pool $\mathcal{S}$. For a pseudonym $id_b^{T\text{-}} \in \mathcal{S}$, the probability $\Pr[id_a^{T+} = id_b^{T\text{-}}|(M_a^{T\text{-}}, id_a^{T\text{-}})]$ that v_a selects $id_b^{T\text{-}}$ as its new pseudonym id_a^{T+} is calculated by the coordinator, as given by

$$\Pr[id_a^{T+} = id_b^{T\text{-}}|(M_a^{T\text{-}}, id_a^{T\text{-}})] = \frac{e^{-\frac{\varepsilon}{2}\mathcal{V}_{b,a}^{T\text{-}}}}{\sum\limits_{id_x^{T\text{-}} \in \mathcal{S}} e^{-\frac{\varepsilon}{2}\mathcal{W}_{x,a}^{T\text{-}}}}, \tag{5.9}$$

where

$$\mathcal{W}_{x,a}^{T\text{-}} = \frac{1}{2}d(\hat{M}_x^{T\text{-}}, M_a^{T\text{-}}) + \frac{1}{2}d(\hat{M}_a^{T\text{-}}, M_a^{T\text{-}}). \tag{5.10}$$

A small $\mathcal{V}_{x,a}^{T\text{-}}$ indicates that $\hat{M}_x$ is close to M_a. $\mathcal{V}_{x,a}^{T\text{-}}$ is negative when $\hat{M}_x$ is closer to M_a than it is to $\hat{M}_a$. Here, v_a has a high probability of selecting $id_b^{T\text{-}}$ as its new pseudonym id_a^{T+} if $\mathcal{V}_{x,a}^{T\text{-}}$ is small. It follows from (5.9) that

$$\mathcal{V}_{x,a}^{T\text{-}} \leq \mathcal{W}_{x,a}^{T\text{-}}, \tag{5.11}$$

which leads to

$$\sum_{id_x^{T\text{-}} \in \mathcal{S}} \Pr[id_a^{T+} = id_x^{T\text{-}} | (M_a^{T\text{-}}, id_a^{T\text{-}})] \leq 1. \tag{5.12}$$

Thus, v_a may not choose any pseudonym when $\sum_{id_x^{T\text{-}} \in \mathcal{S}} \Pr[id_a^{T+} = id_x^{T\text{-}} | (M_a^{T\text{-}}, id_a^{T\text{-}})] < 1$. v_a does not change its pseudonym, i.e., $id_a^{T+} = id_a^{T\text{-}}$, if v_a does not select any id_x in $\mathcal{S}$.

If the vehicles select different pseudonyms, the coordinator returns a pseudonym to each of the vehicles. If multiple vehicles select the same pseudonym $id_x^{T\text{-}}$, the coordinator creates the same number of ring-signature-based pseudonyms using the existing works [409], [410] based on $id_x^{T\text{-}}$ and returns random new ring-signature pseudonyms to the vehicles. Hence, the vehicles can share the same pseudonym $id_x^{T\text{-}}$.

After swapping the pseudonyms at time T, the vehicles use their new pseudonyms to upload their messages with obfuscated driving statuses until the next pseudonym swap. Thus, the probability $\Pr[(\hat{M}_a^{T+}, id_a^{T+}) | (M_a^{T+}, M_a^{T\text{-}}, id_a^{T\text{-}})]$ that v_a with its actual driving statuses $M_a^{T\text{-}}, M_a^{T+}$, and pseudonym $id_a^{T\text{-}}$ generates message $(\hat{M}_a^{T+}, id_a^{T+})$ after swapping pseudonyms can be calculated as

$$\begin{aligned} &\Pr[(\hat{M}_a^{T+}, id_a^{T+}) | (M_a^{T+}, M_a^{T\text{-}}, id_a^{T\text{-}})] \\ =& \Pr[(\hat{M}_a^{T+}, id_a^{T+}) | (M_a^{T+}, id_a^{T+})] \Pr[id_a^{T+} | (M_a^{T\text{-}}, id_a^{T\text{-}})] \\ =& \underbrace{\Pr[\hat{M}_a^{T+} | M_a^{T+}]}_{\text{Obfuscation after swap}} \sum_{\hat{M}_y^{T\text{-}}} \Big(\underbrace{\Pr[id_a^{T+} | (\hat{M}_y^{T\text{-}}, id_a^{T\text{-}})]}_{\text{Swapping pseudonyms}} \underbrace{\Pr[\hat{M}_y^{T\text{-}} | M_a^{T\text{-}}]}_{\text{Obfuscation before swap}} \Big), \end{aligned} \tag{5.13}$$

where $M_a^{T\text{-}}$ and M_a^{T+} are the sequences of the actual driving statuses from v_a before and after swapping the pseudonyms, respectively. $\hat{M}_a^{T\text{-}}$ and $\hat{M}_a^{T+}$ can be any sequences of the obfuscated driving statuses from v_a before and after swapping the pseudonyms, respectively. $\hat{M}_y^{T\text{-}}$ is a possible sequence of the obfuscated driving statuses of v_a before swapping the pseudonyms. In JTOPS, the trajectory obfuscation is independent of the pseudonym swap, but the pseudonym swap relies on the obfuscated statuses.

5.4.3 Privacy Analysis

In this chapter, we first prove that the JTOPS with the spatial-temporal driving statuses satisfies the T-I. Then, we discuss the improvement of the distinguishability upper bounds obtained by jointly considering trajectory obfuscation and pseudonym swapping.

Theorem 4. *For any sequence of obfuscated messages $(\hat{M}_x^{T+}, id_x^{T+})$, the JTOPS mechanism satisfies the $(\varepsilon, \mathbb{R})$-Trajectory-Indistinguishability that any two vehicles v_a and v_b in region $\mathbb{R}$ are $(\varepsilon, \mathbb{R})$-trajectory-indistinguishable.*

Proof. Assuming that the pseudonyms are swapped at time T, we can substitute (5.17), (5.8), and (5.9) into (5.13), and evaluate the distinguishability of v_a and v_b, i.e., the difference of the probabilities that v_a with $(M_a^{T+}, M_a^{T\text{-}}, id_a^{T\text{-}})$ and v_b with $(M_b^{T+}, M_b^{T\text{-}}, id_b^{T\text{-}})$ generate $(\hat{M}_x^{T+}, id_x^{T+})$, as given in (5.11).

$$\frac{\Pr[(\hat{M}_x^{T+},id_x^{T+})|(M_a^{T+},M_a^{T-},id_a^{T-})]}{\Pr[(\hat{M}_x^{T+},id_x^{T+})|(M_b^{T+},M_b^{T-},id_b^{T-})]}=\frac{\frac{e^{-\frac{\varepsilon}{2}d(\hat{M}_x,M_a^{T+})}}{\sum\limits_{\hat{M}_z}e^{-\frac{\varepsilon}{2}d(\hat{M}_z,M_a^{T+})}}\sum\limits_{\hat{M}_y^{T-}}\left(\frac{e^{-\frac{\varepsilon}{2}\mathcal{V}_{x,a}^{T-}}}{\sum\limits_{id_c^{T-}\in\mathcal{S}}e^{-\frac{\varepsilon}{2}\mathcal{W}_{c,a}^{T-}}}\times\frac{e^{-\frac{\varepsilon}{2}d(\hat{M}_y^{T-},M_a^{T-})}}{\sum\limits_{\hat{M}_z}e^{-\frac{\varepsilon}{2}d(\hat{M}_z^{T-},M_a^{T-})}}\right)}{\underbrace{\frac{e^{-\frac{\varepsilon}{2}d(\hat{M}_x,M_b^{T+})}}{\sum\limits_{\hat{M}_z}e^{-\frac{\varepsilon}{2}d(\hat{M}_z,M_b^{T+})}}}_{\text{Obfuscation after swap}}\sum\limits_{\hat{M}_y^{T-}}\left(\underbrace{\frac{e^{-\frac{\varepsilon}{2}\mathcal{V}_{x,b}^{T-}}}{\sum\limits_{id_c^{T-}\in\mathcal{S}}e^{-\frac{\varepsilon}{2}\mathcal{W}_{c,b}^{T-}}}}_{\text{Swapping Pseudonyms}}\times\underbrace{\frac{e^{-\frac{\varepsilon}{2}d(\hat{M}_y^{T-},M_b^{T-})}}{\sum\limits_{\hat{M}_z}e^{-\frac{\varepsilon}{2}d(\hat{M}_z^{T-},M_b^{T-})}}}_{\text{Obfuscation before swap}}\right)},\tag{5.11}$$

In (5.11), $\hat{M}_z$, consisting of the obfuscation candidates $\bar{m}\in\bar{\mathcal{M}}$, is a possible sequence of the driving statuses in region $\mathbb{R}$. $\hat{M}_x^{T-}$ is the sequence of obfuscated driving statuses reported with id_x^{T-}. $(\hat{M}_c^{T-},id_c^{T-})$ is the sequence of the uploaded messages reported by any other vehicle $v_c\in\mathcal{S}$. $\hat{M}_y^{T-}$ is a possible sequence of the driving statuses of v_a and v_b before time T.

To simplify (5.11), we let

$$f(M_j)=\sum_{\hat{M}_z}e^{-\frac{\varepsilon}{2}d(\hat{M}_z,M_j)};\tag{5.12}$$

$$h(M_j)=\sum_{id_c^{T-}\in\mathcal{S}}e^{-\frac{\varepsilon}{2}\mathcal{W}_{c,j}^{T-}};\tag{5.13}$$

$$\mathcal{H}(M_j,\hat{M}_y^{T-})=\frac{1}{h(M_j)}e^{-\frac{\varepsilon}{2}\mathcal{V}_{x,a}^{T-}}\times\frac{1}{f(M_j^{T-})}e^{-\frac{\varepsilon}{2}d(\hat{M}_y^{T-},M_j^{T-})}.\tag{5.14}$$

Thus, (5.11) can be rewritten as

$$\frac{\Pr[(\hat{M}_x^{T+},id_x^{T+})|(M_a^{T+},M_a^{T-},id_a^{T-})]}{\Pr[(\hat{M}_x^{T+},id_x^{T+})|(M_b^{T+},M_b^{T-},id_b^{T-})]}=\frac{f(M_b^{T+})}{f(M_a^{T+})}e^{\frac{\varepsilon}{2}\left(d(\hat{M}_x^{T+},M_b^{T+})-d(\hat{M}_x^{T+},M_a^{T+})\right)}\frac{\sum\limits_{\hat{M}_y^{T-}}\mathcal{H}(M_a,\hat{M}_y^{T-})}{\sum\limits_{\hat{M}_y^{T-}}\mathcal{H}(M_b,\hat{M}_y^{T-})}.\tag{5.15}$$

For any obfuscated driving statuses $(\hat{M}_x^{T-},\hat{M}_x^{T+})$ from v_x, as the driving statuses hold the triangle inequality, we have

$$\mathcal{D}\left(v_b,v_x\right)-\mathcal{D}\left(v_a,v_x\right)\leq\mathcal{D}(v_a,v_b).\tag{5.16}$$

Thus, (5.15) meets the following inequality

$$\frac{\Pr[(\hat{M}_x^{T+},id_x^{T+})|(M_a^{T+},M_a^{T-},id_a^{T-})]}{\Pr[(\hat{M}_x^{T+},id_x^{T+})|(M_b^{T+},M_b^{T-},id_b^{T-})]}\leq\frac{f(M_b^{T+})}{f(M_a^{T+})}\times e^{\frac{\varepsilon}{2}d(M_a^{T+},M_b^{T+})}\times\frac{\sum\limits_{\hat{M}_y^{T-}}\mathcal{H}(M_a,\hat{M}_y^{T-})}{\sum\limits_{\hat{M}_y^{T-}}\mathcal{H}(M_b,\hat{M}_y^{T-})}.\tag{5.17}$$

By employing the triangle inequality, i.e., $d(\hat{M}_z,M_b^{T+})\geq d(\hat{M}_z,M_a^{T+})-d(M_a^{T+},M_b^{T+})$, we have

$$e^{-\frac{\varepsilon}{2}d(\hat{M}_z,M_b^{T+})}\leq e^{-\frac{\varepsilon}{2}\left(d(\hat{M}_z,M_a^{T+})-d(M_a^{T+},M_b^{T+})\right)}.\tag{5.18}$$

Based on (5.18), for all $\hat{M}_z$ consisting of $\bar{m}$, the following holds

$$\sum_{\hat{M}_z}\left(e^{-\frac{\varepsilon}{2}d(\hat{M}_z,M_b^{T+})}-e^{-\frac{\varepsilon}{2}\left(d(\hat{M}_z,M_a^{T+})-d(M_a^{T+},M_b^{T+})\right)}\right)\leq 0,$$

which can be rewritten as

$$\sum\nolimits_{\hat{M}_z} e^{-\frac{\varepsilon}{2}d(\hat{M}_z,M_b^{T+})}-e^{\frac{\varepsilon}{2}d(M_a^{T+},M_b^{T+})}\sum\nolimits_{\hat{M}} e^{-\frac{\varepsilon}{2}d(\hat{M}_z,M_a^{T+})}\leq 0. \tag{5.19}$$

Based on the definition of $f(M_j)$ in (5.12), we have

$$f(M_b^{T+})-e^{\frac{\varepsilon}{2}d(M_a^{T+},M_b^{T+})}f(M_a^{T+})\leq 0, \tag{5.20}$$

which leads to

$$\frac{f(M_b^{T+})}{f(M_a^{T+})}\leq e^{\frac{\varepsilon}{2}d(M_a^{T+},M_b^{T+})}. \tag{5.21}$$

By multiplying $e^{\frac{\varepsilon}{2}d(M_a^{T+},M_b^{T+})}$ on the both sides of (5.21), it follows that

$$\frac{f(M_b^{T+})}{f(M_a^{T+})}\times e^{\frac{\varepsilon}{2}d(M_a^{T+},M_b^{T+})}\leq e^{\varepsilon d(M_a^{T+},M_b^{T+})}. \tag{5.22}$$

Combining (5.17) and (5.22), we have

$$\frac{\Pr[(\hat{M}_x^{T+},id_x^{T+})|(M_a^{T+},M_a^{T\text{-}},id_a^{T\text{-}})]}{\Pr[(\hat{M}_x^{T+},id_x^{T+})|(M_b^{T+},M_b^{T\text{-}},id_b^{T\text{-}})]}\leq e^{\varepsilon d(M_a^{T+},M_b^{T+})}\times\frac{\sum\limits_{\hat{M}_y^{T\text{-}}}\mathcal{H}(M_a,\hat{M}_y^{T\text{-}})}{\sum\limits_{\hat{M}_y^{T\text{-}}}\mathcal{H}(M_b,\hat{M}_y^{T\text{-}})}. \tag{5.23}$$

Based on the definition of $\mathcal{H}(M_j,\hat{M}_y^{T\text{-}})$ in (5.14), we have

$$\frac{\mathcal{H}(M_a,\hat{M}_y^{T\text{-}})}{\mathcal{H}(M_b,\hat{M}_y^{T\text{-}})}=\frac{h(M_b)}{h(M_a)}\times\frac{f(M_b^{T\text{-}})}{f(M_a^{T\text{-}})}\times e^{\frac{\varepsilon}{2}(\mathcal{V}_{x,b}^{T\text{-}}-\mathcal{V}_{x,a}^{T\text{-}})}\times e^{\frac{\varepsilon}{2}\left(d(\hat{M}_y^{T\text{-}},M_b^{T\text{-}})-d(\hat{M}_y^{T\text{-}},M_a^{T\text{-}})\right)}, \tag{5.24}$$

By following the analysis from (5.18) to (5.22), we attain

$$\frac{f(M_b^{T\text{-}})}{f(M_a^{T\text{-}})}\leq e^{\frac{\varepsilon}{2}d(M_a^{T\text{-}},M_b^{T\text{-}})}. \tag{5.25}$$

Based on the triangle inequality of driving statues, we have

$$\frac{e^{\frac{\varepsilon}{2}\left(\frac{1}{2}d(\hat{M}_c^{T\text{-}},M_b^{T\text{-}})-\frac{1}{2}d(\hat{M}_b^{T\text{-}},M_b^{T\text{-}})\right)}}{e^{\frac{\varepsilon}{2}\left(\frac{1}{2}d(\hat{M}_c^{T\text{-}},M_a^{T\text{-}})-\frac{1}{2}d(\hat{M}_a^{T\text{-}},M_a^{T\text{-}})\right)}}\leq e^{\frac{\varepsilon}{2}\left(\frac{1}{2}d(M_a^{T\text{-}},M_b^{T\text{-}})-\frac{1}{2}d(\hat{M}_b^{T\text{-}},M_b^{T\text{-}})+\frac{1}{2}d(\hat{M}_a^{T\text{-}},M_a^{T\text{-}})\right)}. \tag{5.26}$$

As $0\leq d(M_b^{T\text{-}},M_a^{T\text{-}})\leq 1$, we have

$$e^{-\frac{\varepsilon}{2}\left(\frac{1}{2}d(\hat{M}_c^{T\text{-}},M_b^{T\text{-}})+\frac{1}{2}d(\hat{M}_b^{T\text{-}},M_b^{T\text{-}})\right)}\leq e^{-\frac{\varepsilon}{2}\left(\frac{1}{2}d(\hat{M}_c^{T\text{-}},M_a^{T\text{-}})-\frac{1}{2}d(M_b^{T\text{-}},M_a^{T\text{-}})+\frac{1}{2}d(\hat{M}_b^{T\text{-}},M_b^{T\text{-}})\right)}. \tag{5.27}$$

By following the analysis in (5.18) - (5.22), we have

$$\frac{\sum_{id_c^{T\text{-}}\in\mathcal{S}} e^{-\frac{\varepsilon}{2}\left(\frac{1}{2}d(\hat{M}_c^{T\text{-}},M_b^{T\text{-}})+\frac{1}{2}d(\hat{M}_b^{T\text{-}},M_b^{T\text{-}})\right)}}{\sum_{id_c^{T\text{-}}\in\mathcal{S}} e^{-\frac{\varepsilon}{2}\left(\frac{1}{2}d(\hat{M}_c^{T\text{-}},M_a^{T\text{-}})+\frac{1}{2}d(\hat{M}_a^{T\text{-}},M_a^{T\text{-}})\right)}} \leq e^{\frac{\varepsilon}{2}\left(\frac{1}{2}d(M_b^{T\text{-}},M_a^{T\text{-}})+\frac{1}{2}d(\hat{M}_a^{T\text{-}},M_a^{T\text{-}})-\frac{1}{2}d(\hat{M}_b^{T\text{-}},M_b^{T\text{-}})\right)}. \tag{5.28}$$

Based on the definition of $h(M_j)$ in (5.13), it follows:

$$\frac{h(M_b)}{h(M_a)} \leq e^{\frac{\varepsilon}{2}\left(\frac{1}{2}d(M_a^{T\text{-}},M_b^{T\text{-}})-\frac{1}{2}d(\hat{M}_b^{T\text{-}},M_b^{T\text{-}})+\frac{1}{2}d(\hat{M}_a^{T\text{-}},M_a^{T\text{-}})\right)}. \tag{5.29}$$

Combining (5.26) and (5.29), we have

$$\frac{h(M_b)}{h(M_a)} \times \frac{e^{\frac{\varepsilon}{2}\mathcal{V}_{x,b}^{T\text{-}}}}{e^{\frac{\varepsilon}{2}\mathcal{V}_{x,a}^{T\text{-}}}} \leq e^{\frac{\varepsilon}{2}\left(d(M_a^{T\text{-}},M_b^{T\text{-}})-d(\hat{M}_b^{T\text{-}},M_b^{T\text{-}})+d(\hat{M}_a^{T\text{-}},M_a^{T\text{-}})\right)}, \tag{5.30}$$

which leads to

$$\frac{\mathcal{H}(M_a^{T\text{-}},\hat{M}_y^{T\text{-}})}{\mathcal{H}(M_b^{T\text{-}},\hat{M}_y^{T\text{-}})} \leq e^{\varepsilon d(M_a^{T\text{-}},M_b^{T\text{-}})}. \tag{5.31}$$

As $\hat{M}_y^{T\text{-}}$ can be any possible obfuscated driving status, the following holds

$$\frac{\sum\limits_{\hat{M}_y^{T\text{-}}} \mathcal{H}(M_a^{T\text{-}},\hat{M}_y^{T\text{-}})}{\sum\limits_{\hat{M}_y^{T\text{-}}} \mathcal{H}(M_b^{T\text{-}},\hat{M}_y^{T\text{-}})} \leq e^{\varepsilon d(M_a^{T\text{-}},M_b^{T\text{-}})}. \tag{5.32}$$

By substituting (5.32) into (5.23), we finally have

$$\frac{\Pr[(\hat{M}_x^{T+},id_x^{T+})|(M_a^{T+},M_a^{T\text{-}},id_a^{T\text{-}})]}{\Pr[(\hat{M}_x^{T+},id_x^{T+})|(M_b^{T+},M_b^{T\text{-}},id_b^{T\text{-}})]} \leq e^{\varepsilon d(M_a^{T\text{-}},M_b^{T\text{-}})} e^{\varepsilon d(M_a^{T+},M_b^{T+})} = e^{\varepsilon\mathcal{D}(v_a,v_b)}, \tag{5.33}$$

which satisfies (5.16). In other words, for any sequence of uploaded messages $(\hat{M}_x^{T+},id_x^{T+})$, vehicles v_a and v_b in region $\mathbb{R}$ are $(\varepsilon,\mathbb{R})$-trajectory-indistinguishable. Hence, the proposed JTOPS mechanism can protect privacy under the external GPA.

□

We find that pseudonym swapping can be considered a differential privacy process.

Corollary 3. *Pseudonym swapping can be treated as a differential privacy process.*

Proof. Let id' be any pseudonym after pseudonym swapping, and $\{id_1, id_2, \cdots, id_n\}$ be the identity set that can be swapped to id'. For any id_i and id_j $(i, j \in [1, n])$, the following holds,

$$\begin{cases} 0 < \Pr[id'|id_i] \leq 1, \\ 0 < \Pr[id'|id_j] \leq 1, \end{cases}$$

Let $\max(\frac{\Pr[id'|id_i]}{\Pr[id'|id_j]}) = x$. According to the definition of $\Pr[id'|id_i]$ and $\Pr[id'|id_j]$, it is easy to prove $x \geq 1$.

Thus we have

$$\frac{1}{x} \leq \frac{\Pr[id'|id_i]}{\Pr[id'|id_j]} \leq x. \tag{5.34}$$

Let $\varepsilon = \ln x$. The above equation can be transformed to

$$e^{-\varepsilon} \leq \frac{\Pr[id'|id_i]}{\Pr[id'|id_j]} \leq e^{\varepsilon}, \tag{5.35}$$

which is an expression of differential privacy.

□

We next prove the superiority of the JTOPS, compared to the pseudonym-swapping-only or trajectory obfuscation-only approach.

Corollary 4. *Applying pseudonym swapping and trajectory obfuscation jointly can provide higher location privacy-preserving capability than using either of the pseudonym swapping and trajectory obfuscation.*

Proof. When separately using the pseudonym swapping and trajectory obfuscation, the pseudonyms are swapped based on the actual driving statuses. For any two vehicles v_a and v_b at time T, the mechanism, which separately swaps the pseudonyms and obfuscates the driving statuses, can be written as

$$\frac{\Pr[(\hat{M}_a^{T\text{-}}, id_a^{T+})|(M_a^{T\text{-}}, id_a^{T\text{-}})]}{\Pr[(\hat{M}_b^{T\text{-}}, id_b^{T+})|(M_b^{T\text{-}}, id_b^{T\text{-}})]} = \underbrace{\frac{\Pr[\hat{M}_a^{T\text{-}}|M_a^{T\text{-}}]}{\Pr[\hat{M}_b^{T\text{-}}|M_b^{T\text{-}}]}}_{\text{Obfuscation before swap}} \times \underbrace{\frac{\Pr[id_a^{T+}|id_a^{T\text{-}}]}{\Pr[id_b^{T+}|id_b^{T\text{-}}]}}_{\text{Swapping pseudonym}}. \tag{5.36}$$

Assuming that the separate use of the pseudonym swapping and trajectory obfuscation is based on DP, according to the definition of $d(M_a^{T+}, v_b^{T+})$ and corollary 1, we have

$$\begin{cases} \dfrac{\Pr[\hat{M}_a^{T\text{-}}|M_a^{T\text{-}}]}{\Pr[\hat{M}_b^{T\text{-}}|M_b^{T\text{-}}]} \leq e^{\varepsilon d(M_a^{T\text{-}}, M_b^{T\text{-}})}; \\ \dfrac{\Pr[id_a^{T+}|id_a^{T\text{-}}]}{\Pr[id_b^{T+}|id_b^{T\text{-}}]} \leq e^{\varepsilon'}, \end{cases}$$

where ε' is the privacy budget of the pseudonym-swapping process. The above equations lead to

$$\frac{\Pr[(\hat{M}_a^{T+}, id_a^{T+})|(M_a^{T+}, M_a^{T-}, id_a^{T-})]}{\Pr[(\hat{M}_b^{T+}, id_b^{T+})|(M_b^{T+}, M_b^{T-}, id_b^{T-})]} \leq e^{\varepsilon d(M_a^{T-}, M_b^{T-})+\varepsilon'}. \tag{5.37}$$

In the same case, according to the definition of $\mathcal{H}(M_j, \hat{M}_y^{T-})$ in (5.14) and (5.18) - (5.31), the JTOPS mechanism can be written as

$$\frac{\Pr[(\hat{M}_a^{T-}, id_a^{T+})|(M_a^{T-}, id_a^{T-})]}{\Pr[(\hat{M}_b^{T-}, id_b^{T+})|(M_b^{T-}, id_b^{T-})]} = \frac{\mathcal{H}(M_a^{T-}, \hat{M}_y^{T-})}{\mathcal{H}(M_b^{T-}, \hat{M}_y^{T-})} \leq e^{\varepsilon d(M_a^{T-}, M_b^{T-})}. \tag{5.38}$$

By comparing the right-hand sides of (5.37) and (5.38), jointly utilizing the trajectory obfuscation and pseudonym swapping has a lower upper bound of indistinguishability, indicating the JTOPS achieves higher privacy protection than using trajectory obfuscation or pseudonym swapping only. □

The JTOPS swaps the pseudonyms based on the obfuscated driving statuses of the vehicles, which does not increase the upper bound of the indistinguishability.

5.4.4 Under Collusion Attack

An adversary can potentially collude with the coordinator that is aware of the pseudonym swapping results. Then, the old and new pseudonyms can be precisely linked. In this chapter, we analyze how the JTOPS can limit the coordinator's knowledge so as to protect the location privacy of the vehicles. Compared with the existing pseudonym-based mechanisms [384], [401], our mechanism can still protect the location privacy in the presence of an untrusted or collusive coordinator.

If an adversary can precisely link the new and old pseudonyms, the Estimation Probability (EP) that the adversary can accurately estimate the actual driving statuses of the vehicle is defined as follows.

Definition 9 (EP)**.** *Given a set of obfuscated driving statuses $\hat{M} = \{\hat{m}_1, \cdots, \hat{m}_i, \cdots\}$ from vehicle v with the actual driving statuses $M = \{m_1, \cdots, m_i, \cdots\}$, the EP of the vehicle i.e., $\theta(v)$, is given by*

$$\theta(v) = \frac{1}{k} \sum_{i=1}^{k} \Pr[m_i|\hat{m}_i], \tag{5.39}$$

where $\hat{m}_i$ is the obfuscated driving status of vehicle v based on the i-th actual m_i. $\Pr[m_i|\hat{m}_i]$ is based on Bayes' theorem [15]. A low EP indicates a high privacy-preserving capability.

As the pseudonym swapping results are known by the adversary, we introduce our previous work RN-I [15] to show the JTOPS can still provide the location privacy. By extending the metrics, the RN-I in this chapter is defined as follows.

Definition 10 (RN-I)**.** *A mechanism satisfies* $(\varepsilon, \mathbb{R})$*-RN-I if and only if, with obfuscated driving statuses* $\hat{M}_a$*, any actual driving status pair* (M_{a_1}, M_{a_2}) *of* v_a *yield*

$$\frac{\Pr[\hat{M}_a|M_{a_1}]}{\Pr[\hat{M}_a|M_{a_2}]} \leq e^{\varepsilon d(M_{a_1}, M_{a_2})}. \tag{5.40}$$

Theorem 5. *If an adversary colludes with the coordinator, the JTOPS mechanism can protect the privacy of vehicles, such that, for the obfuscated driving statuses* $\hat{M}_a$*, any set of driving statuses* M_{a_1} *and* M_{a_2} *of vehicle* v_a *satisfies RN-I.*

Proof. We assume that pseudonym swapping results are exposed. Thus, the JTOPS would be written as

$$\frac{\Pr[\hat{M}_a|M_{a_1}]}{\Pr[\hat{M}_a|M_{a_2}]}, \tag{5.41}$$

where $\hat{M}_a$ is the reported trajectory of vehicle v_a.

By substituting (5.17) into (5.41), the latter can be written as

$$\frac{\Pr[\hat{M}_a|M_{a_1}]}{\Pr[\hat{M}_a|M_{a_2}]} = \frac{\frac{e^{-\frac{\varepsilon}{2}d(M_{a_1}, \hat{M}_a)}}{\sum_{\hat{M}_{a_x}} e^{-\frac{\varepsilon}{2}d(M_{a_1}, \hat{M}_{a_x})}}}{\frac{e^{-\frac{\varepsilon}{2}d(M_{a_2}, \hat{M}_a)}}{\sum_{\hat{M}_{a_x}} e^{-\frac{\varepsilon}{2}d(M_{a_2}, \hat{M}_{a_x})}}}, \tag{5.42}$$

where $\hat{M}_{a_x}$ is any sequence of obfuscated driving statuses that can be generated by vehicle v_a.

By following the analysis from (5.18) to (5.22), we have

$$d(\hat{M}_a, M_{a_2}) - d(\hat{M}_a, M_{a_1}) \leq d(M_{a_1}, M_{a_2}). \tag{5.43}$$

Let $g(M_{a_i}) = \sum_{\hat{M}_x} e^{-\frac{\varepsilon}{2}d(M_{a_i}, \hat{M}_x)}$, we have

$$\frac{\Pr[\hat{M}_a|M_{a_1}]}{\Pr[\hat{M}_a|M_{a_2}]} = \frac{g(M_{a_2})}{g(M_{a_1})} e^{\frac{\varepsilon}{2}\left(d(M_{a_2}, \hat{M}_a) - d(M_{a_1}, \hat{M}_a)\right)} \leq \frac{g(M_{a_2})}{g(M_{a_1})} e^{\frac{\varepsilon}{2}d(M_{a_1}, M_{a_2})}. \tag{5.44}$$

By employing the triangle inequality, we have

$$e^{-\frac{\varepsilon}{2}d(M_{a_2}, \hat{M}_x)} \leq e^{-\frac{\varepsilon}{2}\left(d(M_{a_1}, \hat{M}_x) - d(M_{a_1}, M_{a_2})\right)}. \tag{5.45}$$

Following the analysis of (5.18)–(5.22), we have

$$\frac{g(M_{a_2})}{g(M_{a_1})} \leq e^{\frac{\varepsilon}{2}d(M_{a_1}, M_{a_2})}. \tag{5.46}$$

By substituting (5.46) into (5.44), the latter can be rewritten as

$$\frac{\Pr[\hat{M}_a|M_{a_1}]}{\Pr[\hat{M}_a|M_{a_2}]} \leq e^{\varepsilon d(M_{a_1}, M_{a_2})}, \tag{5.47}$$

which satisfies the RN-I. In other words, for any sequence $\hat{M}_a$ of obfuscated driving statuses from vehicle v_a, any two historical driving status sequences M_{a_1} and M_{a_2} in the prior knowledge are $(\varepsilon, \mathbb{R})$-Road Network-indistinguishable. Hence, the proposed JTOPS mechanism can resist internal GPAs. □

5.5 Experimental Results

In this chapter, we compare the JTOPS with the existing hybrid mechanism [75], which is also based on pseudonyms and obfuscation in the same scenario as considered in this chapter. We also compare the trajectory obfuscation process and pseudonym swapping process of the JTOPS with the two latest mechanisms developed in [15] and [384].

The experiments are conducted with two real-world road networks[2] and the T-Drive trajectory (10,357 drivers in Beijing, China) [342]. The two real-world road networks are extracted from the Open Street Map[3]. We utilize the trajectories of vehicles, which have at least ten trajectory points in the experimental road networks. We set $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0.25$. In the experiments, we average the ASR (5.11), DU (5.12), PU (5.13), and EP (5.39) of multiple vehicles by running the mechanisms locally on the vehicles for 50 times.

5.5.1 Adversary's Success Rate and Estimation Probability

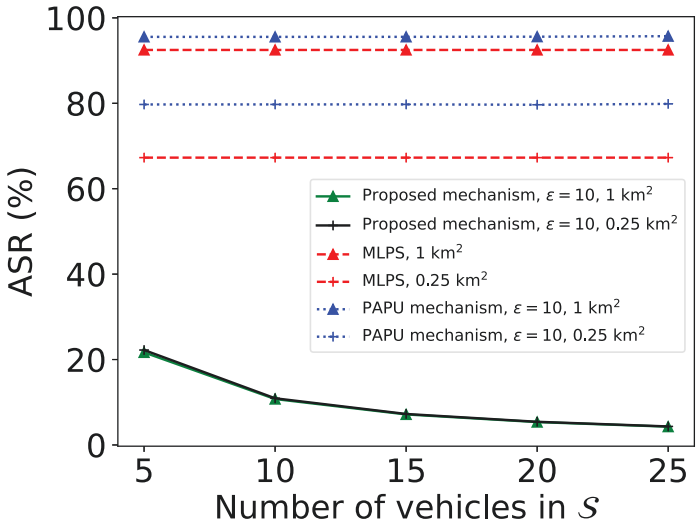
The comparison of the ASR (5.11) among the JTOPS, Multilevel Location Privacy Scheme (MLPS) [75], and Pseudonym Swap With Provable Unlinkability (PAPU) [384] mechanism is shown in Fig. 5.2(a). The ASRs of PAPU and MLPS are higher than 80%, and do not change with ε and the number of vehicles in $\mathcal{S}$. The reason is that the two mechanisms require the vehicles to report actual driving statuses. Adversaries can infer the actual identity of each vehicle and ignore pseudonym swaps by matching the obtained driving statuses with their prior knowledge of the vehicles. The ASR of the PAPU mechanism is higher than that of the MLPS in both large and small regions, as the MLPM randomly selects cooperative vehicles and the PAPU selects candidates with pseudonym swapping probabilities. Compared with the two mechanisms, the proposed mechanism achieves a low ASR in both two road networks. The JTOPS allows the vehicles to upload obfuscated driving statuses so that the adversary cannot gain any useful information by eavesdropping. Even if the adversary has the prior knowledge of the vehicles, the driving statuses can still be protected by the proposed JTOPS.

The ASRs of the JTOPS with different numbers of vehicles, ε, and region sizes are shown in Fig. 5.2(b). The ASR increases with ε. It is evident that the JTOPS can provide better privacy protection with a smaller ε. In the large region size, the ASR of the JTOPS is approximately 21% and 20.6% when ε is 10 and 5 vehicles. When 25 vehicles swap their pseudonyms, the ASR decreases to 5% in both $\varepsilon = 5$ and $\varepsilon = 10$. The adversary cannot gain useful knowledge by analyzing the obtained driving statuses of the vehicles as the uploaded driving statuses are obfuscated. As a result, the adversary cannot match their prior knowledge, and it can only randomly guess the actual identity of the vehicle. By randomly guessing the actual identity, the number of vehicles in $\mathcal{S}$ has a stronger impact on the ASR than ϵ.

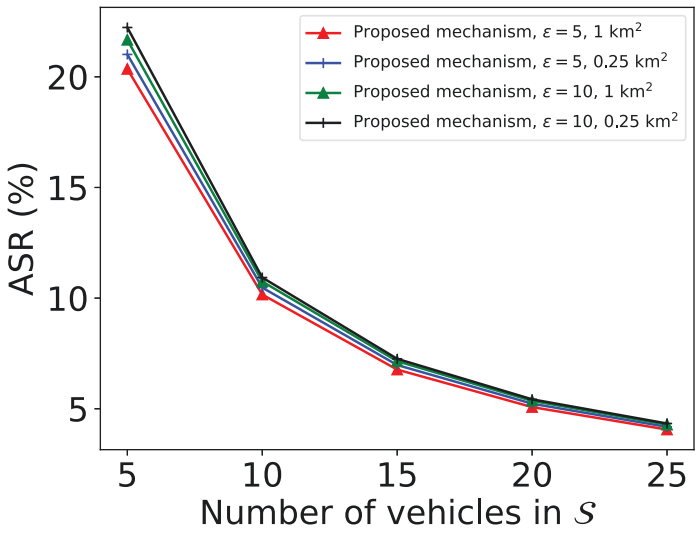
The comparison of the EP (5.39) between the JTOPS and Personalized Location Privacy-Preserving mechanism (PLPP) [15] is shown in Fig. 5.2(c). The PLPP mechanism takes the route distance as the only metric, while the JTOPS extends the metric to the location distance, direction, speed, and driving time of the vehicles. As shown in Fig. 5.2(c), the proposed

[2]There are 3,246 trajectory points of 119 vehicles, 34 connections and 63 roads within the small region size (116.3595°$\leq$ longitude $\leq$ 116.3645°and 39.9135°$\leq$ latitude $\leq$ 39.9085°). There are 20,745 trajectory points of 558 vehicles, 181 connections, and 386 roads within the large region size (116.357°$\leq$ longitude $\leq$ 116.367°and 39.906°$\leq$ latitude $\leq$ 39.916°).

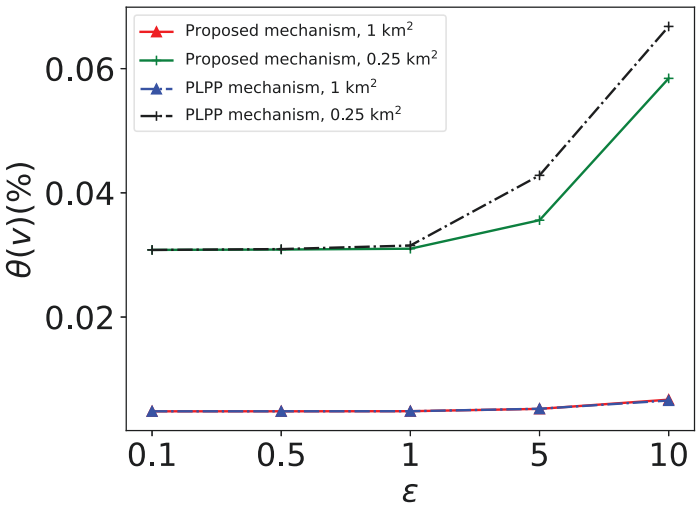
[3]Open Street Map is an open-source database of the world's geographic map. `https://www.openstreetmap.org/`



(a) The comparison of Adversary's Success Rate (ASR) among our JTOPS, MLPS [75], and the PAPU mechanism [384].



(b) The ASR of JTOPS.



(c) The comparison of EP between the JTOPS and PLPP mechanism [15].

Figure 5.2: The comparison of privacy protection and data utility.

Table 5.2: Performance in Different Privacy Requirements (MLPM does not balance privacy and utility).

Trade-off			Mechanism			
			Proposed JTOPS	PAPU [384]	PLPP [15]	MLPM [75]
High Privacy ($\varepsilon = 0.1$)	ADU	1 km^2	68.8%	N/A	**69.1%**	68.2%
		0.25 km^2	**61.3%**	N/A	60.6%	**75.9%**
	PU	1 km^2	**77.2%**	72.5%	N/A	N/A
		0.25 km^2	**80.4%**	68.9%	N/A	N/A
	ASR (10 vehicles in $\mathcal{S}$)	1 km^2	**10.0%**	93.2%	N/A	91.6%
		0.25 km^2	**10.2%**	80.1%	N/A	77.2%
High Utility ($\varepsilon = 100$)	ADU	1 km^2	**77.7%**	N/A	72.3%	68.2%
		0.25 km^2	75.8%	N/A	72.4%	**75.9%**
	PU	1 km^2	**79.3%**	73.9%	N/A	N/A
		0.25 km^2	**79.1%**	70.4%	N/A	N/A
	ASR (10 vehicles in $\mathcal{S}$)	1 km^2	**11.1%**	93.2%	N/A	91.6%
		0.25 km^2	**11.9%**	80.1%	N/A	72.2%

mechanism provides a higher privacy-preserving capability than the PLPP mechanism in the small region. The distance between the two mechanisms is significant under a high ε value. In the large region, the two mechanisms have similar estimation probabilities of less than 0.01%.

5.5.2 Data Utility and Pseudonym Utility

We proceed to compare the trajectory obfuscation process and the pseudonym swapping process of the JTOPS with the PLPP and PAPU mechanisms in Figs. 5.3. As shown in Fig. 5.3, the trajectory obfuscation process of the JTOPS, which considers multiple driving statuses, achieves a higher data utility than the PLPP mechanism. The improvement of data utility is significant under a high ε because the mechanism has a high probability of selecting the candidate close to the actual vehicle. The ADU (5.12) of the JTOPS is higher in data utility in the large regions than it is in the small regions. This is because the obfuscation candidate set is fine-grained in the large region, as the densities of the two regions are similar. Thus, the vehicles have more choices, which are close to their actual driving statuses, in the large region than they have in the small region. The difference of the DU between the JTOPS and PLPP is smaller in the small regions than it is in the large regions, because the driving statuses between the vehicles and candidates are close in a small region.

As shown in Fig. 5.3, the pseudonyms have a high data utility by using the JTOPS, that the driving statuses reported with the pseudonym before the swapping process are similar to those of the pseudonym after the swapping process. The reason is that the actual driving statuses of the vehicles in road networks have high randomness, indicating the distance in the actual driving statuses among various vehicles could be significant. The proposed mechanism utilizes

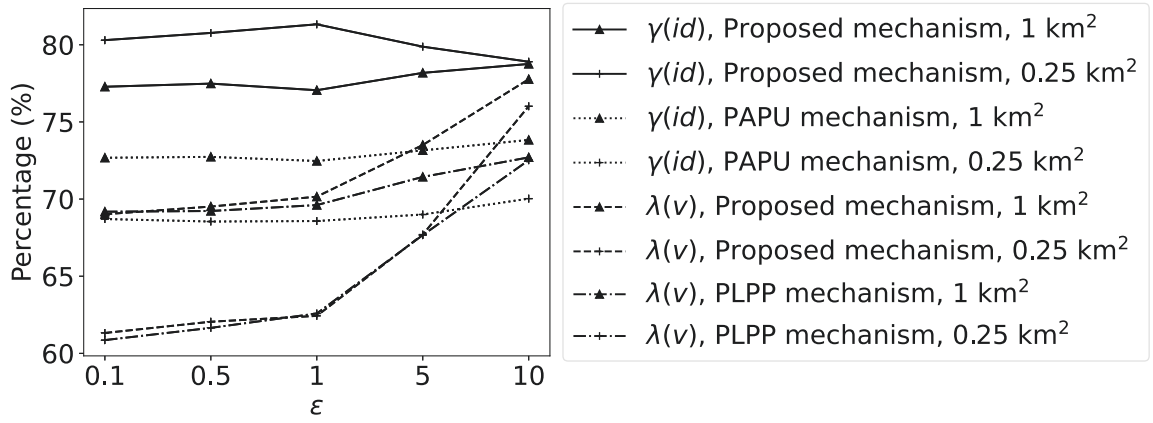


Figure 5.3: The comparison of data and pseudonym utility.

obfuscated driving statuses that are selected from the obfuscated candidates. Thus, the distance in the obfuscated driving statuses is stable in fixed road networks. As the large region and the small region have a similar road density, the driving statuses of the candidates in the small region are closer to each other than those in the large region. For the PAPU mechanism, the driving status granularity of the vehicles in the small region is coarser than those in the large region, so the PU of the PAPU mechanism in the small region is less than it is in the large region.

5.5.3 Performance in Different Privacy Requirements

The comparisons of ADU, PU, and ASR under different privacy requirements among the JTOPS, PAPU [384], PLPP [15], MLPM [75] are shown in Table 5.2. We follow the privacy setting of [15], where $\varepsilon = 0.1$ represents that the drivers need a high level of privacy protection and $\varepsilon = 10$ represents that drivers need a high level of data and pseudonym utility. A high ADU and PU correspond to a high utility of data and pseudonyms, while a low ASR means high privacy protection.

In the ASR comparison, we assume there are ten vehicles in the pseudonym swapping process. Compared with PLPP, the ASRs of PAPU and MLPM are high, and the values are around 80% in the small region and more than 90% in the large regions. This is because the two mechanisms do not protect the uploaded messages that the adversary can accurately match the pseudonyms with the actual identities based on its prior knowledge. The ASRs of PAPU and MLPM do not change with ε because ε does not influence the contents of the uploaded messages. The JTOPS achieves a significant location privacy-preserving capability that its ASR values are almost 10% in each case.

In the utility comparison, the ADUs of the JTOPS and PLPP are significantly higher in the high-utility requirement case than they are in the high-privacy requirement case. This is because the two mechanisms have a high probability of selecting candidates who are similar to the actual driving statuses. The ADU of MLPM is lower than those of the two mechanisms because MLPM selects cooperators randomly. The ADU of PAPU is not considered in this comparison, as PAPU does not protect actual driving statuses. The PUs of the JTOPS and PAPU are similar, but the PU of JTOPS is a bit higher than that of PAPU. One reason is that the JTOPS addresses the conflict of pseudonym swapping, which is overlooked by PAPU. Another reason is that the pseudonym swapping probability functions of the two mechanisms are different, where the

JTOPS allows the vehicles to select a similar candidate with a high probability. Nevertheless, the JTOPS still satisfies the concept of DP, which is explained as follows.

Let the actual driving status sequence of vehicle v is $M = \{m_1, \cdots, m_l\}$, and the obfuscation candidate set is $\bar{\mathcal{M}} = \{\bar{m}_1, \cdots, \bar{m}_n\}$. The obfuscated driving statuses sequence of v is $\hat{M} = \{\hat{m}_1, \cdots, \hat{m}_l\}$. As each obfuscated driving status $\hat{m}$ is independently selected, according to (5.17), we have

$$\Pr[\hat{M}|M] = \prod_{m_i \in M} \frac{e^{-\frac{\varepsilon}{2} d(m_i, \hat{m}_i)}}{\sum_{\bar{m}_x \in \bar{\mathcal{M}}} e^{-\frac{\varepsilon}{2} d(m_i, \bar{m}_x)}}. \tag{5.48}$$

For any $\hat{m}_i \in \hat{M}$, $\hat{m}$ is selected from $\bar{\mathcal{M}}$, and each $\bar{m} \in \bar{\mathcal{M}}$ can be selected multiple times, so that the denominator in (5.48) can be transformed into

$$\prod_{m_i \in M} \sum\nolimits_{\bar{m}_x \in \bar{\mathcal{M}}} e^{-\frac{\varepsilon}{2} d(m_i, \bar{m}_x)} = \sum\nolimits_{\hat{M}_x} e^{-\frac{\varepsilon}{2} d(M, \hat{M}_x)}, \tag{5.49}$$

where $\hat{M}_x$ is the possible sequence of driving statuses that consists of multiple $\hat{m}_x$.

Based on the definition of $d(M_1, M_2)$ in (5.1), we have

$$\prod_{m_i \in M} e^{-\frac{\varepsilon}{2} d(m_i, \hat{m}_i)} = e^{-\frac{\varepsilon}{2} d(M, \hat{M})}. \tag{5.50}$$

Thus, (5.48) can be rewritten as

$$\Pr[\hat{M}|M] = \frac{e^{-\frac{\varepsilon}{2} d(M, \hat{M})}}{\sum_{\hat{M}_x} e^{-\frac{\varepsilon}{2} d(M, \hat{M}_x)}}. \tag{5.51}$$

5.6 Conclusion

In this chapter, we defined a unified privacy-preserving measure, i.e., T-I, for mechanisms that use trajectory obfuscation and pseudonym swapping and satisfy ε-DP. Using the T-I, we proposed a new JTOPS, which is the first attempt to jointly employ trajectory obfuscation and pseudonym swapping in road networks to protect both the identity and location privacy of vehicles. The JTOPS was proved to achieve the T-I, and validated with comprehensive experiments. The JTOPS can provide privacy protection in the case of GPAs with the full prior knowledge or collusion attacks.

Chapter 6

Conclusions and Future Work

6.1 Lessons Learnt and Open Challenges for LPPM in Future Vehicular Networks

This section starts with illustrating the localization techniques in future vehicular networks, which are compared with their counterparts in current versions. By discussing the new requirements of location privacy and data utility introduced by the novel localization, we present the limitations of the existing LPPMs in future vehicular networks. The potential location privacy issues under the future cross-layer tracking techniques introduced by new communication technologies are illustrated, and we offer solutions to these issues by improving the existing LPPMs.

6.1.1 Advancement of Localization vs. Location Privacy

Localization technologies in future 5G/6G-enabled vehicular networks can be classified as basic localization, cooperative localization, machine learning-based localization, driver tracking, and multipath exploiting localization [411], as shown in Table 6.1. The differences in localization techniques between the future vehicular networks and the current ones are also shown in Table 6.1. We compare the BS density, recourse cost, and localization precision of the localization techniques in Table 6.1. In future vehicular networks, localization, sensing, and communication will coexist, sharing the same time-frequency-spatial resources [412], [413]. The localization technologies in future vehicular networks are as follows [412].

Traditional Localization

Traditional localization technologies mainly focus on geometric calculation with augmented assistance, under the assumption that the positions of the Base Stations (BSs) are knowable and the positions of User Equipment (UEs) are geometric constrained [414]. For example, the position of UEs can be calculated if the distances and physical angular orientations between BSs and UEs are known. Compared with the localization precision of the existing cellular radio-based tracking, the future basic localization could have high precision due to the high data rates and scalability. The frequent communication and the high BSs density also expose a granularity trajectory. The traditional localization techniques can be classified into distance measurement, angle measurement, area measurement, and hop-count measurement [415], as follows.

Table 6.1: Difference of localization techniques in current vehicular networks and future vehicular networks.

Localization techniques	**Features of future vehicular networks**	**Corresponding techniques in current vehicular networks**	**Difference**
Traditional localization	High data rates High scalability	Cellular radio-based tracking	More BS and UE More frequent signal transmission More accurate localization
Cooperative localization	Ultra-low latency High scalability	Sensing infrastructure-based tracking Cellular radio-based tracking Upper-layer message-based tracking	More sensors More frequent signal transmission Closer BS and UE
AI-based localization	High data rates High scalability	Sensing infrastructure-based tracking Optical vision-based tracking Upper-layer message-based tracking	More sensors More detailed dataset Massive devices
Channel charting	High scalability	Cellular radio-based tracking Sensing infrastructure-based tracking Upper-layer message-based tracking	Considering channel state information
Driver tracking localization	High data rates Ultra-low latency	Sensing infrastructure-based tracking Upper-layer message-based tracking	More sensors 3D vehicular network
Multipath exploiting localization	Ultra-low latency High scalability	Cellular radio-based tracking	More sensors

SLAM	High data rates Ultra-low latency High scalability	Vehicle driving log-based tracking Upper layer message-based tracking Cellular radio-based tracking	More sensors Considering time-varying states More fixed landmarks
Context-aware localization	High data rates Ultra-low latency	Vehicle driving log-based tracking Upper-layer message-based tracking	Intelligent prediction Massive personal information

• *Distance Measurement:* Distance measurement calculates the positions with the obtained distance-related measurements. The distance-related measurements leverage transmission time (e.g., time difference of arrival), received signal strength indicator, and connectivity condition [416].

• *Angle Measurement:* Angle measurement employs trigonometry and geometric calculation, which requires the distance and angular information between the target and BSs [417].

• *Area Measurement:* Area measurement localization technology uses the center intersection of all overlapping coverage regions as the estimation position [418]. The more restricted zones are obtained, the higher precision can be achieved [419].

• *Hop-count Measurement:* This kind of localization can be realized by analyzing the number of hops [420], [421].

Cooperative Localization

Cooperative localization has been employed in current vehicular networks, as the vehicles and other devices can communicate with others directly through D2D communication protocols [422]. Cooperative localization allows UEs to measure the distance and angular information on each D2D link [423]. UEs are typically closer to each other than to the BSs in D2D communication, which has higher signal-to-noise ratio [424]. This method provides a high position precision for cooperative localization. Due to the ultra-low latency and high scalability, massive nodes in future vehicular networks would be connected in the network, and the distance between UEs would be closer than that in current version. Hence, the precision of cooperative localization would be significantly more accurate in future vehicular networks than in current version.

AI-based Localization

AI-based localization is a data-centric technique [411]. The localization algorithms utilize machine learning to generate a fingerprint database of the environment [425]. The fingerprinting database contains the channel parameters (e.g., channel state information) that are measured at the known locations (reference points) [426]. Then, the localization algorithms can compare the information of the environment surrounding vehicles to estimate the drivers' positions. In future vehicular networks, the dataset for machine learning could be more detailed due to the

features like hierarchical coexistence, flexible storage, and flexible processing.

Table 6.2: Localization technologies in future vehicular networks.

Localization	BS Density	Cost	Precision
Traditional	High	Low [427]	Low [428]
Cooperative	Low	Low [429]	Medium [430]
AI-based	Medium	High [431]	High [431]
Channel Charting	Depends	Depends	Depends
Driver Tracking	Medium	Low [432]	Medium [432]
Multipath Exploiting	Low	Medium [433]	High [434]
SLAM	Low	Depends	High [435]
Context-aware	High	Medium [436]	High [437]

Channel Charting

Channel charting localization employs AI-based algorithms [438]. It generates a virtual map based on the gained channel state information. The drivers can be located and tracked on the virtual map. Although the map cannot provide the drivers' actual locations, it offers a real-time pseudo-location as reference [412].

Driver tracking localization

Driver tracking localization continuously infers the driver's position to smooth out the estimation errors [411]. This localization technique can predict the drivers' trajectories with the vehicle sensor data as follows.

• *Passive Sensing:* Passive sensing in future vehicular networks is also known as passive radar or passive coherent location [412]. Passive sensing locates the target by receiving and processing the energy reflected by the targets [439]. In future sensing infrastructure-based vehicular networks, it is almost impossible to avoid disclosing location privacy by sensor signals.

• *Active Sensing:* Active sensing (radar) localization is employed in various scenarios, e.g., adaptive cruise control and cross-traffic alerts [440]. In future vehicular networks, sensors could share the data related to the location information to support the high-precision localization applications. The sensor localization will provide high-precision distance in 3D vehicular networks [412].

Multipath-based Localization

Multipath-based localization discloses the drivers' positions by combining multipath components and environment geometry [441]. The multipath components can be seen as mirror images of a physical environment [442]. Using the road network environment knowledge, multipath-based localization can be launched with the information of the radar signal path. Thus, the multipath-based localization can exploit additional position-related data in radio signals, even if the vehicles are not in the line of sight [443]. As the vehicles in future vehicular networks would equip with multiple sensors, the radar-like signals from the sensors could be utilized in multipath-based localization.

Simultaneous Localization And Mapping (SLAM)

High data rates, ultra-low latency, and high scalability of the future vehicular networks introduce SLAM for the vehicles. SLAM localization aims to estimate the state of the vehicles and the landmarks [444]. The vehicles and other mobile devices in SLAM consider vehicles with time-varying states [435]. In future vehicular networks, the SLAM can be classified into vision-based SLAM and radar-based SLAM as follows.

• *Vision-based SLAM*: It uses image sensors (e.g., cameras) to detect landmarks whose states are fixed or changing slowly.

• *Radar-based SLAM*: It requires laser sensors that can provide higher accuracy than vision-based sensors. 3D lidar point clouds would be widely used in future autonomous vehicular networks.

The entities in future vehicular networks can collect location information with radar-like signals from other vehicles, fixed infrastructures, and sensors [445].

Context-aware Localization

The future 5G/6G-enabled vehicular networks allow intelligent data transmission and multi-modal localization prediction by combining highly personal information with public data [437]. The drivers can select and change communication channels and technologies according to their locations and contexts [436]. Thus, future vehicular networks could utilize the message context in channels for localization.

6.1.2 Limitations and Opportunities of LPPMs for Future Upper Layer Location Privacy Attacks

With the development of vehicular networks, there would be new challenges for the existing LPPMs. The existing LPPMs mainly focus on upper-layer message-based tracking that the adversaries with other tracking techniques (i.e., sensing infrastructure-based, optical vision-based, vehicle driving log-based, and cellular radio-based) cannot be defended in the existing vehicular networks. Nevertheless, as shown in Table 6.3, the existing LPPMs could not provide acceptable location privacy protection under the upper-layer message-based attack scenario in future vehicular networks.

User-side LPPM

Most of the existing user-side LPPMs cannot suit the low latency of the future vehicular networks, i.e., pass-and-run, certificate, and secure computation. Although data perturbation can satisfy the latency requirement, the location data protected by data perturbation cannot provide high-precision LBSs for high-precision localization applications.

• *Pass-and-Run:* The communication delay in the pass-and-run is high since the vehicles apply LBSs by routing through other nodes. Identity verification is also a challenge for the pass-and-run-based vehicles because it would lead to high communication delays for legal cooperator authentication [446]. The frequent communication among multiple vehicles could also increase the precision of signal-based localization, i.e., basic localization and cellular radio-based localization.

Table 6.3: Challenge of location privacy in future vehicular networks.

Techniques	**New challenges**	**LPPMs that are available**	**LPPMs that need to be improved**
THz	Small coverage area Spectrum penetration power Centimeter-level precision via APs	Pass-and-run Certificates for Privacy Secure computation Homomorphic encryption Private information retrieval Searchable encryption Secure communication	Data perturbation Statistical disclosure control Trusted third party
VLC	Signal scatter Observation signal Effective transmission requirement	Certificates for Privacy Secure communication Trusted third party	Pass-and-run Data perturbation Secure computation Statistical disclosure control Private information retrieval Homomorphic encryption Searchable encryption
mmWave	Exchange CSI frequently Low latency requirement Eavesdrop channel easily	Certificates for Privacy Secure computation Data perturbation Statistical disclosure control Homomorphic encryption Private information retrieval Searchable encryption Trusted third party Secure communication	Pass-and-run

Sub-6 GHz	Flexible antenna design Fake BS and malicious devices In-vehicle tracking	Pass-and-run Certificates for Privacy Secure communication Trusted third party	Secure computation Data perturbation Statistical disclosure control Homomorphic encryption Private information retrieval Searchable encryption
Satellite Communication	Long propagation delay Difficult to allocate MIMO Communication resource allocation	Certificates for Privacy Data perturbation Statistical disclosure Secure communication Trusted third party	Pass-and-run Secure computation Homomorphic encryption Private information retrieval Searchable encryption
QC	Quantum channel Quantum computer	Certificates for Privacy Secure computation Data perturbation Statistical disclosure control Secure communication Trusted third party	Pass-and-run Homomorphic encryption Searchable encryption Private information retrieval

The transmission delay caused by routing can be decreased by sensors that are fully used in future vehicular networks. With the sensors, the drivers can have full knowledge of the road network and communication environment to select the channels and routers (i.e., vehicles and sensors) [447]. The trust issues of the pass-and-run can be solved by trust management, i.e., authentication and certificate. The advanced blockchain technologies [448] would be utilized in future vehicular networks, which can also be employed for trust management.

• *Certificates for Privacy:* Certificate-based LPPMs achieve authentication by verifying the certifications and signatures [182]. There are two limitations of the certificate-based LPPMs in future vehicular networks: requiring a trusted authority and storage consumption of certificate management [380]. The delay caused by the authentication process might be unacceptable for communication consumption [449]. The massive nodes, i.e., vehicles and sensors, would increase the storage consumption, which would be challenging to manage certificates.

The features of blockchain, i.e., immutable, secure, and consensus, satisfy the authentication requirement [17]. Therefore, the blockchain platform could be utilized as a support for certificates [284]. The information collected by sensors could also be referred to for authentication.

• *Secure Computation:* The major limitation of secure computation is computation delay, although it satisfies the requirement of flexible storage and processing. Even in current vehicular networks, the computation consumption of the existing secure computation is still a problem [5].

• *Data Perturbation:* Data perturbation is flexible for storage and processing [15]. The communication and calculation delays of the data perturbation are also low. Nevertheless, the future vehicular networks require highly accurate location data, i.e., centimeter-level or millimeter-level, but the existing data-perturbation-based LPPMs generally generate meter-level precision [450]. If the data perturbation provides centimeter-level or millimeter-level precision, the location privacy-preserving capability of the LPPMs could be extremely low [451].

The scenario for data perturbation in future vehicular networks would be limited. Hence, it should be combined with other LPPMs and use adaptive noise to satisfy the future scenario.

Server-side LPPM

The low latency requirement also challenges the existing server-side LPPMs, as the process at the server side could bring computation delay.

• *Statistical Disclosure Control:* The statistical disclosure control is achieved by using LPPMs, like data perturbation, secure computation, and anonymity on the server side [222]. Therefore, the limitations of the corresponding LPPMs on the user side also exist on the server side. As data transmission in future vehicular networks would be ultra frequent and the amount of data would be huge, computing consumption, storage consumption, and data management would be great challenges for statistical disclosure control.

The data management and storage consumption of statistical disclosure control could be optimized by utilizing blockchain-based techniques, but the computing consumption would be difficult to decrease [227]. AI-based LPPMs can be launched to achieve adaptive protection and efficient analysis to simplify the process [51]. Trust management could also be employed in statistical disclosure control to avoid costing computational resources on trusted entities.

• *Homomorphic Encryption:* Homomorphic encryption is one of the popular, secure computations. The server-side homomorphic encryption is also limited by its computation delay, which

can lead to a high computational and communication consumption [250]. It is necessary to simplify the process of homomorphic encryption for computational consumption reduction.

• *Private Information Retrieval:* The computational consumption of the PIR cloud be high, which is difficult to be allocated in current vehicular networks. For the future vehicular networks with low-latency requirements, the current version of PIR would also be impossible to be launched in practice.

• *Searchable Encryption:* The LBSs in future vehicular networks require high-precision location data, but the existing SE schemes would return results with errors to the drivers. The existing works allow the drivers to use SE and other LPPMs together to improve the accuracy of results. Nevertheless, the computational consumption could also be increased when SE schemes are combined with certificated-based or other tools.

Information from sensors in future vehicular networks could be referred to correct the result, as massive sensors would be allocated. Also, AI-based LPPMs can be introduced to optimize the result according to the driver's historical information, which would be fine-grained [452].

User-server-interface LPPM

The user-server-interface LPPMs would attract more researchers' attention in future vehicular networks than in the current version because future vehicular networks would ask the vehicles to share data through frequent V2X communication. Protecting location privacy in channels would be an approach to defend basic localization, channel charting localization, and driver tracking localization.

• *Secure Communication:* The communication channel protocols of future vehicular networks would be improved to be different from the current protocols. As the protocols of current vehicular networks have not been completely developed, we do not consider the protocols in this part. For the end-to-end encrypted services of secure communication, it is seldom used in current vehicular networks [273]. The computational consumption of data encryption and decryption could be high, increasing the latency of communication.

• *Trusted Third Party (TTP):* TTP in current vehicular networks is an ideal scenario which is difficult to achieve privacy, as the adversaries can hijack the TTP even if the TTP is controlled by government[1]. In future vehicular networks, there will be massive sensors and vehicles which can generate a huge amount of data. A hijacked TTP could lead to serious privacy disclosure. The trust issues of the massive sensors and vehicles also reduce the efficiency of TTP. To solve these limitations, TTP should be combined with other LPPMs, such as encryption, blockchain, AI-based LPPMs, anonymity, and data perturbation.

6.1.3 Location Privacy Challenges and Emerging Wireless Technologies

Different from the existing vehicular networks, the future version transmits data through new communication mediums. The new communication mediums improve the efficiency of vehicular networks but bring new cross-layer location privacy issues.

[1] bleepingcomputer.com/news/security/hacker-claims-to-have-stolen-data-on-1-billion-chinese-citizens/

• *Sub-6 GHz:* The Sub-6 GHz has a spectrum range from 0.45 GHz to 6 GHz and wavelength from 0.15 cm to 2 cm [115]. The wide coverage capability and low cost have been extensively investigated in the existing works to support V2X communications [453].

The sub-6 GHz techniques introduce new challenges to the existing physical-layer LPPMs because of the flexible antenna design [454]. However, the flexible antenna design allows the pass-and-run to transform data effectively. By using fake BSs and malicious devices, the adversaries can obtain distance and angular data by monitoring multiple sub-6 GHz links, with which the adversaries can calculate the positions of vehicles in high precision [455]. The above two issues could be solved by using the certificate and a trusted third party, which aims to detect illegal entities for authentication. Nevertheless, data perturbation and statistical disclosure control increase the difficulty of illegal entity detection.

Malicious applications can be installed on mobile devices or vehicles for tracking the vehicles. The Pegasus developed in Israel can be installed on Android and iOS mobile phones to monitor drivers' location data. The in-vehicle communication via the sub-6 GHz channel can expose the vehicles' location data, as mobile phones have similar location data to the vehicles if the drivers carry mobile phones when driving [142]. To protect location privacy in such a scenario, the driver can hide its semantic information in the location data and block the adversaries' eavesdropping links. Hence, the user-server-interface LPPMs and server-side LPPMs would attract attention to the sub-6 GHz techniques when combined with the AI and blockchain techniques.

• *mmWave:* The mmWave has a spectrum range from 30 GHz to 100 GHz and a wavelength from 1 nm to 10 mm [456]. Due to the beam-based directional transmissions and the utilization of the huge spectrum, mmWave can reach a high multi-gigabit speed and communicate in all weather [457].

In future vehicular networks, the frequent changing of the network topology and channel requires nodes (e.g., vehicles, RSUs, and sensors) to exchange CSI with a low delay, which increases communication consumption [458]. As the V2X communication needs a low transmission delay, the LPPMs in future vehicular networks should be highly efficient. Pass-and-run will increase the transmission delay, so it cannot satisfy the low latency requirement of mmWave.

The adversaries can eavesdrop on the mmWave to obtain and infer the transmitted location data [459]. However, the Doppler shift of signal transmission, which is unavoidable in future mmWave-based vehicular networks, can be combined with the data perturbation and anonymity to protect location privacy [460]. Then, the adversaries cannot obtain the real information by eavesdropping. Secure computation, homomorphic encryption, private information retrieval, and searchable encryption could be used to prevent the adversaries from obtaining the useful information by eavesdropping because they do not expose the private information in the channel.

• *Terahertz (THz) communication:* The THz has a spectrum range from 100 GHz to 10 THz and a wavelength from 30 μm to 3000 μm [461]. Nevertheless, the communication region of the THz is small, i.e., less than 50 m [462]. The frameworks of the THz-based vehicular networks have been designed in the existing studies, e.g., [463]–[465]. To overcome the limitation of small coverage regions and improve the utilization of THz bands, the design in [463] used an SDN controller to select an optimal route between the source and the destination.

As shown in Table 6.3, the existing LPPMs would face new challenges with THz communication.

The pass-and-run mechanism could be improved with the THz-based routing, but the trust issues would be more serious in future vehicular networks than that in current version. Hence, the certificate-based LPPMs would be popular in future THz-band communication. Although the low coverage area and penetration power of THz spectrum increase the privacy level of the communication medium, THz-band communication requires the vehicles' private information, e.g., whereabouts, driving status, road network information, and traffic condition [463]. The potential optimization for the above two drawbacks could be overcome by combining the trust management mechanisms and blockchain or AI-based techniques. As the coverage area of the THz is small, the data perturbation and statistical disclosure control cannot provide expected location privacy protection with a small number of perturbation candidates. The numerous legal sensors in future networks could also be a good choice to route the data.

The Access Points (APs) for THz-band communication, which are a kind of trusted third party, can achieve centimeter-level precision of localization when tracking the vehicle's mobility to overcome the limitations of the existing vehicular networks (e.g., frequent network disconnection and volatility of connection) [466]. Nevertheless, the existing DP-based user-side LPPMs cannot satisfy the high-precision localization in such a scenario. Thus, the DP-based LPPMs on the server side have more widespread application prospects than they are on the user side. In future vehicular networks, the user-side DP-based LPPMs should focus on more driving states rather than only using location data to achieve high data utility [467].

• *Visible Light Communication (VLC):* Lighting devices in vehicular networks, such as traffic signals, roadside lights, and vehicle lights, can be employed for VLC to send data with Light-Emitting Diode (LED) installations [468]. VLC has a spectrum range from 400 THz to 800 THz and a wavelength from 380 μm to 780 μm [469]. VLC assists the communication for V2X with low energy consumption and high efficiency [470].

Due to the inevitable scatter issues, VLC outside is not as safe as it is inside [471], and it can be eavesdropped on by the adversaries. For outside VLC for V2X, the scattered signals can be received by the adversaries, i.e., eavesdropper [472], even though the techniques like non-orthogonal multiple access [473] and pseudo surface waves [474] can minimize the attenuation effect. Therefore, the pass-and-run in VLC will increase the risk of eavesdropping. Data perturbation and statistical disclosure control cannot protect location data when the source of the signal can be physically observed. As the signal has been observed on the user side, the server-side LPPMs cannot provide location privacy protection.

As the privacy risk introduced by the use of VLC is related to its physical channel characteristics, the certification for authentication and the user-server-interface LPPMs could outperform other existing LPPMs. The certification-based methods can achieve high location privacy protection in short-range VLC by preventing unauthorized receivers from joining VLC, as the scattered signals in shot-range VLC can be ignored [471]. When VLC is allocated for remote communication, the physical layer LPPMs, e.g., cooperative jamming [475] and Channel State Information (CSI) estimation [472], can be used to prevent the eavesdropper from obtaining the transmitted private information. Cooperative jamming is a positive LPPM that aims to block the eavesdropper's channel by allowing multiple legal nodes to route the signal [476]. The CSI estimation is a passive LPPM that asks the legal nodes to infer the CSI of legal and eavesdropping links with which the legal nodes can select different devices to send signals [472].

• *Quantum Communication (QC):* The QC utilizes the quantum states of lights, which achieves

secure communication by using microscopic particles to carry quantum information [477].

Quantum Key Distribution (QKD) has been integrated with existing classical optical networks in QC to achieve cost-efficient and secure communication [478]. Tradition cryptography that relies on mathematical computations is challenged by the development of quantum computing, making QKD important in QC [479]. QKD is based on Heisenberg's uncertainty principle and quantum no-cloning theorem enabling detection of eavesdropping on the key distribution [480]. However, QKD cannot protect location privacy in vehicular networks, as the laser for QC can be used to infer the direction of the source and destination vehicles. To address this limitation, quantum homomorphic encryption [481], quantum searchable encryption [482], and quantum private information retrieval [483] have been developed for secure QC based on QKD. These techniques enhance the security of QC by enabling computation on encrypted data without decryption, secure search of encrypted data, and retrieval of data from a database without revealing the query contents, respectively.

To enable scalable quantum communication with massive nodes, researchers have been exploring quantum teleportation (QT) in long-range communication scenarios. QT allows for the sender to divide information into traditional and quantum channels and transfer the message to the receiver through both channels, with the message at the sender side being destroyed during transmission. This method has received significant attention due to its potential for log-range communication [484], [485]. However, the use of QT in vehicular networks presents a challenge to existing routing-based PPMs, such as pass-and-run, as they become ineffective in this scenario [485].

Quantum Identity Authentication (QIA) and Quantum Signature (QS) are introduced to realize trust management for secure and reliable QC. The trust management in QC can be classified into objective trust and subjective trust, as follows,

- *Objective Trust:* The trust management is based on certifiable evidence, e.g., certification.
- *Subjective Trust:* The trust management is based on a kind of group, which has specific characteristics or behaviors, e.g., TTP.

Hence, the certification-based LPPM and statistical disclosure control can be improved to achieve a secure QC in vehicular networks.

• *Reconfigurable Intelligent Surface (RIS):* Comprising passive reflect arrays and control elements, RIS provides a programmable wireless environment for the vehicle-road-human integrated network, which exploits the advantages of the novel electromagnetic wave manipulation technique, i.e., metamaterials [486]. As there are no specific or complicated requirements for the installation location of the metamaterials, RIS can be easily installed, e.g., on the building facades and billboards, which are close to the roads for communication support [487]. The metamaterials would have low computational and energy consumption [488]. Hence, RIS is conducive to future large-scale deployment, which can also conform to the concept of green communication and sustainable development. The support of full-duplex and full-band communication in RIS provides comprehensive coverage and diverse options for vehicular communications [489]. Furthermore, Multi-Input Multi-Output (MIMO) technology could be applied to RIS to utilize the large surface area of RIS for better antenna deployment, which shows great potential in future vehicular networks [490]. However, the mobility and personalization of the drivers in vehicular networks have been overlooked by the existing RIS design [491]. The dynamic network topology and disconnection of vehicular networks lead to a low utility

of RIS in vehicular networks, as the RIS cannot gain knowledge from the frequently changed environment or the limited feedback from the drivers [491].

6.1.4 Challenges Arising from Networks Convergence

In the integrated vehicular networks, the vehicles can exchange information and communicate with other vehicles, roadside infrastructure, and pedestrians automatically through real-time V2X communications [492]. With the support of V2X technology, traffic information (e.g., vehicle status, live road conditions, and pedestrian information) enable the formation of the integrated vehicle-road-human network [115]. The components with 5G/6G characteristics in the integrated vehicle-road-human network are illustrated as follows.

• *Integrated Satellite and Terrestrial Network (ISTN):* The satellite communication, which is vulnerable to changeable weather, can provide future vehicular 3D services while cooperating with the existing ground vehicular networks [115].

Satellite communication has the limitations such as long propagation delay, difficulty in allocating massive Multiple-Input Multiple-Output (MIMO) networks, and communication resource allocation [453]. Therefore, the pass-and-run is difficult to be used in satellite communication because of its transmission delay. Other existing LPPMs, e.g., homomorphic encryption and secure communication, which have high computational consumption and communication delay, should be optimized to improve their efficiency. To overcome the above limitations, the recourse consumption methods could be run on trusted third parties. The data perturbation and statistical disclosure control can protect location privacy effectively so that they can outperform their counterparts in satellite communications.

• *Human Interaction:* The rapid development of mobile devices carried by pedestrians has brought powerful communication and data processing capabilities to enable information interaction between the pedestrians and vehicles [493]. Communication between the vehicles and humans (including the pedestrians, vehicle drivers and passengers) will be more frequent and efficient in future vehicular networks, which can be utilized to improve the vehicular networks [494]. By sharing data actively and passively with the pedestrians, the vehicles can overcome the limitations of blind spots in vehicular sensor cameras. And thus, the vehicles can obtain services like accurate situation reporting and early warnings that can reduce the traffic accidents and protect lives and property [495]. With the mobile devices equipped by the drivers and passengers, the vehicles can communicate with the base stations efficiently and can autonomously complete the required tasks. For instance, the vehicles without GPS can also be accurately located with the smart devices of passengers [496]. Frequent communication between the vehicles and pedestrians increases the risk of cross-layer attacks. The encounter information can be combined to infer the trajectory of the target vehicle. Trust management-based LPPMs can be employed to prevent the adversaries who can localize the drivers based on the in-vehicle driving log and sensing information will be more serious.

6.2 Summary of Outcomes

In this thesis, we started by reviewing the existing localization techniques and LPPMs, where their advantages and limitations are discussed. To balance location privacy protection and data utility in real-world road networks, we proposed a personalized location privacy-preserving

mechanism with a novel Road-Indistinguishability Theorem. Based on the RN-I, we then proposed two mechanisms: obfuscated vehicular trajectory detection and cloak region obfuscation. The former allows LBS providers to detect malicious vehicles without breaching the location privacy protection, while the latter extends the parameters of the proposed RN-I. By using anonymity and differential privacy, we proposed a novel Joint Trajectory Obfuscation and Pseudonym Swapping mechanism, which is proven to combine two differential privacy processing without introducing the additivity composition theorem of ε-DP.

6.3 Recommendations & Future Work

In future work, we will find the correlation between differential privacy and anonymity. We have found that t-closeness is equivalent to differential privacy, and ε-DP can achieve e^{ε}-closeness anonymity using only DP-based obfuscated results. The draft of this is about to be submitted. The discovery will be extended to other scenarios, such as split learning and quantum communication.

Chapter 7

Publication List

7.1 First-author Paper

1. **Ma, Baihe**, et al. Personalized Location Privacy With Road Network-Indistinguishability. IEEE Transactions on Intelligent Transportation Systems 23.11 (2022): 20860-20872.

2. **Ma, Baihe**, et al. New Cloaking Region Obfuscation for Road Network-Indistinguishability and Location Privacy. Proceedings of the 25th International Symposium on Research in Attacks, Intrusions and Defenses. 2022.

3. **Ma, Baihe**, et al. Vehicle Trajectory Obfuscation and Detection. Cybersecurity for Smart Cities. Springer, Cham, 2023. 121-134.

7.2 Co-author Papers

1. Jiang, Y.* and **Ma, Baihe***, et al. A Secure Aggregation for Federated Learning on Long-Tailed Data. Accepted by 2022 Conference on Neural Information Processing Systems (NeurIPS) Workshop.

2. Zhao, Y.* and **Ma, Baihe***, et al. Trajectory Obfuscation and Detection in Internet-of-vehicles. 2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design (CSCWD). IEEE, 2022.

3. Wang, Z.* and **Ma, Baihe***, et al. Differential Preserving in XGBoost Model for Encrypted Traffic Classification. 2022 International Conference on Networking and Network Applications (NaNA). IEEE, 2022.

4. Shi, K.* and **Ma, Baihe***, et al. Layered classification method for darknet traffic based on Weighted K-NN. 2022 International Conference on Networking and Network Applications (NaNA). IEEE, 2022.

5. Guo, X.* and **Ma, Baihe***, et al. Fine-Grained Defense Methods in Encrypted Traffic Inspection.“ The 15th China Computer Networks and Information Security Conference & 2022 International Cyberspace Security Academic Summit Forum. 2022.

6. Guo, X.* and **Ma, Baihe***, et al. Fine-Grained Defense Methods in Federated Encrypted Traffic Inspection”. Accepted by Journal of Xidian University.

7. Lin, X., **Ma, Baihe**, Wang, X., He, Y., Liu, R.P. and Ni, W., 2022, May. Multi-layer reverse engineering system for vehicular controller area network messages. In 2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design (CSCWD) (pp. 1185-1190). IEEE.

8. Liang, L., Lin, X., **Ma, Baihe**, Wang, X., He, Y., Liu, R.P. and Ni, W., 2022, December. Leveraging Byte-Level Features for LSTM-based Anomaly Detection in Controller Area Networks. In GLOBECOM 2022-2022 IEEE Global Communications Conference (pp. 4903-4908). IEEE.

9. Shi, K., Zeng, Y., **Ma, Baihe**, Liu, Z., and Ma, J., 2023, MT-CNN: A Classification Method of Encrypted Traffic Based on Semi-Supervised Learning. Accepted 2023 IEEE Global Communications Conference: Communication & Information Systems Security. IEEE.

10. Wang, Z. and **Ma, Baihe** et al. 2023, DLPriv: Deep Learning based Dynamic Location Privacy Mechanism for LBS in Internet-of-Vehicles. Accepted by 2023 International Conference on Networking and Network Applications (NaNA2023).

7.3 Under Review

7.3.1 First-author Paper

1. **Ma, Baihe** et al., "Unlinkable Obfuscation Mechanism for Road Network-Indistinguishability and Location Privacy." submitted to IEEE Transactions on Information Forensics and Security, second round review.

2. **Ma, Baihe** et al., "Location Privacy Threats and Protections in Future Vehicular Networks: A Comprehensive Review." submitted to IEEE Communications Surveys and Tutorials, second round review.

7.3.2 Co-author Paper

1. Jiang, Y.* and **Ma, Baihe*** et al., "Blockchained Federated Learning for Internet of Things: A Comprehensive Survey." submitted to ACM Computing Surveys.

2. Tan, S., Wang, C., Shen, Y., Qian, C., **Ma, Baihe**, and Gao, J., "Dynamic Model for Umbrella Antenna Pointing Performance and Sensitivity Considering Inter-rib Tension Rope Stiffness Uncertainty." submitted to Journal of Aerospace Engineering.

3. Jiang, Y.* and **Ma, Baihe*** et al., "Unmasking Privacy Vulnerabilities in Federated Learning: A New Attack Vector through Data Attribute Correlation in Differential Privacy." submitted to IEEE Transactions on Neural Networks and Learning Systems.

4. Lin, X.* and **Ma, Baihe*** et al., "CAN-Trace Attack: Exploit CAN Messages to Uncover Driving Trajectories." submitted to IEEE Transactions on Intelligent Transportation Systems.

Bibliography

[1] H. Hasrouny, A. E. Samhat, C. Bassil, and A. Laouiti, "VANET security challenges and solutions: A survey," *Veh. Commun.*, vol. 7, pp. 7–20, 2017.

[2] V. Primault, A. Boutet, S. B. Mokhtar, and L. Brunie, "The long road to computational location privacy: A survey," *IEEE Commun. Surv. Tutor.*, vol. 21, no. 3, pp. 2772–2793, 2018.

[3] F. Farouk, Y. Alkady, and R. Rizk, "Efficient privacy-preserving scheme for location based services in VANET system," *IEEE Access*, vol. 8, pp. 60 101–60 116, 2020.

[4] Z. Tan, C. Wang, C. Yan, M. Zhou, and C. Jiang, "Protecting privacy of location-based services in road networks," *IEEE Trans. Intell. Transp. Syst.*, pp. 1–14, 2020. DOI: 10.1109/TITS.2020.2992232.

[5] J. Zhou, Z. Cao, Z. Qin, X. Dong, and K. Ren, "LPPA: Lightweight privacy-preserving authentication from efficient multi-key secure outsourced computation for location-based services in VANETs," *IEEE Trans. Inf. Forens. Secur.*, vol. 15, pp. 420–434, 2019.

[6] B. Li, R. Liang, D. Zhu, W. Chen, and Q. Lin, "Blockchain-based trust management model for location privacy preserving in VANET," *IEEE Trans. Intell. Transp. Syst.*, 2020.

[7] Z. Lu, G. Qu, and Z. Liu, "A survey on recent advances in vehicular network security, trust, and privacy," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 2, pp. 760–776, 2018.

[8] H. Zheng and H. Hu, "Missile: A system of mobile inertial sensor-based sensitive indoor location eavesdropping," *IEEE Trans. Inf. Forens. Secur.*, vol. 15, pp. 3137–3151, 2019.

[9] F. Altché and A. de La Fortelle, "An LSTM network for highway trajectory prediction," in *2017 IEEE 20th international conference on intelligent transportation systems (ITSC)*, IEEE, 2017, pp. 353–359.

[10] A. Sarker, C. Qiu, H. Shen, H. Uehara, and K. Zheng, "Brake data-based location tracking in usage-based automotive insurance programs," in *2020 19th ACM/IEEE Int. Conf. Inf. Proc. Sens. Netw. (IPSN)*, IEEE, 2020, pp. 229–240.

[11] B. Liu, W. Zhou, T. Zhu, L. Gao, and Y. Xiang, "Location privacy and its applications: A systematic study," *IEEE Access*, vol. 6, pp. 17 606–17 624, 2018.

[12] Y. Wang, S. Wang, S. Zhang, and H. Cen, "An edge-assisted data distribution method for vehicular network services," *IEEE Access*, vol. 7, pp. 147 713–147 720, 2019.

[13] H. U. Mustakim, "5G vehicular network for smart vehicles in smart city: A review," *J. Comput. Electron. Telecommun.*, vol. 1, no. 1, pp. 13–18, 2020.

[14] N. Malik, P. Nanda, A. Arora, X. He, and D. Puthal, "Blockchain based secured identity authentication and expeditious revocation framework for vehicular networks," in *Proc. 17th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun. / 12th IEEE Int. Conf. Big Data Sci. Eng. (TrustCom/BigDataSE)*, IEEE, 2018, pp. 674–679.

[15] B. Ma, X. Wang, W. Ni, and R. P. Liu, "Personalized location privacy with road network-indistinguishability," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 11, 2022.

[16] S. Tiwari, S. Kaushik, P. Jagwani, and S. Tiwari, "A survey on LBS: System architecture, trends and broad research areas," in *Proc. Int. Workshop Databases Networked Inf. Syst.*, Springer, 2011, pp. 223–241.

[17] B. Luo, X. Li, J. Weng, J. Guo, and J. Ma, "Blockchain enabled trust-based location privacy protection scheme in VANET," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 2034–2048, 2019.

[18] A. M. Basahel, A. A. Abi Sen, M. Yamin, and S. Alqahtani, "Bartering method for improving privacy of LBS," *Int. J. Comput. Sci. Netw. Secur.*, vol. 19, no. 2, pp. 207–213, 2019.

[19] L. Zhu, C. Zhang, C. Xu, X. Du, N. Guizani, and K. Sharif, "Traffic monitoring in self-organizing VANETs: A privacy-preserving mechanism for speed collection and analysis," *IEEE Wirel.*, vol. 26, no. 6, pp. 18–23, 2019.

[20] D. Suo, J. Moore, M. Boesch, K. Post, and S. E. Sarma, "Location-based schemes for mitigating cyber threats on connected and automated vehicles: A survey and design framework," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 4, pp. 2919–2937, 2022. DOI: 10.1109/TITS.2020.3038755.

[21] A. P. Mdee, M. M. Saad, M. Khan, M. T. R. Khan, and D. Kim, "Impacts of location-privacy preserving schemes on vehicular applications," *Veh. Commun.*, p. 100 499, 2022.

[22] C. Zhang, L. Zhu, C. Xu, *et al.*, "TPPR: A trust-based and privacy-preserving platoon recommendation scheme in VANET," *IEEE Trans. Serv. Comput.*, vol. 15, no. 2, pp. 806–818, 2022. DOI: 10.1109/TSC.2019.2961992.

[23] A. Leander, "Parsing pegasus: An infrastructural approach to the relationship between technology and swiss security politics," *Swiss Political Sci. Rev.*, vol. 27, no. 1, pp. 205–213, 2021.

[24] I. Butun and M. Gidlund, "Location privacy assured Internet of Things.," in *Proc. ICISSP*, vol. 19, Prague, Czech Republic, 2019, pp. 1–8.

[25] R. Gupta and U. P. Rao, "Achieving location privacy through CAST in location based services," *J. Commun. Netw.*, vol. 19, no. 3, pp. 239–249, 2017.

[26] B. Liu, W. Zhou, T. Zhu, L. Gao, and Y. Xiang, "Location privacy and its applications: A systematic study," *IEEE Access*, vol. 6, pp. 17 606–17 624, 2018.

[27] M. Duckham and L. Kulik, "Location privacy and location-aware computing," in *Dynamic & Mobile GIS: Investigating Change in Space and Time*, vol. 3, CRC press Boca Raton, Fla, USA, 2006, pp. 35–51.

[28] W. Han and Y. Xiao, "Privacy preservation for V2G networks in smart grid: A survey," *Comput. Commun.*, vol. 91, pp. 17–28, 2016.

[29] M. A. Al-shareeda, M. Anbar, I. H. Hasbullah, and S. Manickam, "Survey of authentication and privacy schemes in vehicular ad hoc networks," *IEEE Sens. J.*, vol. 21, no. 2, pp. 2422–2433, 2021.

[30] M. A. Al-Shareeda, M. Anbar, S. Manickam, and A. A. Yassin, "VPPCS: VANET-based privacy-preserving communication scheme," *IEEE Access*, vol. 8, pp. 150 914–150 928, 2020.

[31] P. M. Rao and B. Deebak, "Security and privacy issues in smart cities/industries: Technologies, applications, and challenges," *J. Ambient Intell. Humaniz. Comput.*, pp. 1–37, 2022.

[32] A. Pfitzmann and M. Hansen, *A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management*, 2010.

[33] I. Almalkawi, J. Raed, A. Alsarhan, A. Abdallah, and E. Abdallah, "A novel and efficient priority-based cross-layer contextual unobservability scheme against global attacks for WMSNs," *Int. J. Online Eng.*, 2022.

[34] Y. Yu, J. Zhang, and K. B. Letaief, "Joint subcarrier and CPU time allocation for mobile edge computing," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM) 2016*, IEEE, 2016, pp. 1–6.

[35] X. Zha, X. Wang, W. Ni, *et al.*, "Analytic model on data security in vanets," in *Proc. 17th Int. Symp. Commun. Inf. Technol. (ISCIT)*, IEEE, 2017, pp. 1–6.

[36] M. N. Tahir, M. Katz, and U. Rashid, "Analysis of VANET wireless networking technologies in realistic environments," in *Proc. IEEE Radio Wireless Symp.*, IEEE, 2021, pp. 123–125.

[37] A. Boukerche and N. Aljeri, "Design guidelines for topology management in software-defined vehicular networks," *IEEE Netw.*, vol. 35, no. 2, pp. 120–126, 2021.

[38] Q. Cui, X. Hu, W. Ni, *et al.*, "Vehicular mobility patterns and their applications to Internet-of-Vehicles: A comprehensive survey," *Sci. China Inf. Sci.*, vol. 65, no. 11, pp. 1–42, 2022.

[39] M. Dixit, R. Kumar, and A. K. Sagar, "VANET: Architectures, research issues, routing protocols, and its applications," in *Proc. Int. Conf. Comput. Commun. Autom. (ICCCA) 2016*, IEEE, 2016, pp. 555–561.

[40] A. Alnasser, H. Sun, and J. Jiang, "QoS-balancing algorithm for optimal relay selection in heterogeneous vehicular networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 8223–8233, 2022. DOI: 10.1109/TITS.2021.3076901.

[41] A. M. Vegni, C. B. Iglesias, and V. Loscri, "MOVES: A memory-based vehicular social forwarding technique," *Comput. Netw.*, vol. 197, p. 108 324, 2021.

[42] N. K. Chaubey, "Security analysis of Vehicular Ad hoc Networks (VANETs): A comprehensive study," *Int. J. Secur. its Appl.*, vol. 10, no. 5, pp. 261–274, 2016.

[43] M. Dibaei, X. Zheng, Y. Xia, *et al.*, "Investigating the prospect of leveraging blockchain and machine learning to secure vehicular networks: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 2, pp. 683–700, 2022. DOI: 10.1109/TITS.2020.3019101.

[44] L. Chao, C. Wu, T. Yoshinaga, W. Bao, and Y. Ji, "A brief review of multipath TCP for vehicular networks," *Sensors*, vol. 21, no. 8, p. 2793, 2021.

[45] M. A. Mujahid, K. A. Bakar, T. S. Darwish, and F. T. Zuhra, "Cluster-based location service schemes in VANETs: Current state, challenges and future directions," *Telecommun. Syst.*, vol. 76, no. 3, pp. 471–489, 2021.

[46] K. N. Qureshi, A. Alhudhaif, A. A. Shah, S. Majeed, and G. Jeon, "Trust and priority-based drone assisted routing and mobility and service-oriented solution for the Internet of Vehicles networks," *J. Inf. Secur. Appl.*, vol. 59, p. 102 864, 2021.

[47] R. Soua, E. Kalogeiton, G. Manzo, *et al.*, "SDN coordination for CCN and FC content dissemination in VANETs," in *Ad Hoc Netw.* Springer, 2017, pp. 221–233.

[48] R. G. Engoulou, M. Bellaı oche, S. Pierre, and A. Quintero, "VANET security surveys," *Comput. Commun.*, vol. 44, pp. 1–13, 2014.

[49] M. Faraji-Biregani and R. Fotohi, "Secure communication between UAVs using a method based on smart agents in unmanned aerial vehicles," *J. Supercomput.*, vol. 77, no. 5, pp. 5076–5103, 2021.

[50] G. Karmakar, A. Chowdhury, R. Das, J. Kamruzzaman, and S. Islam, "Assessing trust level of a driverless car using deep learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 4457–4466, 2021. DOI: 10.1109/TITS.2021.3059261.

[51] J. Posner, L. Tseng, M. Aloqaily, and Y. Jararweh, "Federated learning in vehicular networks: Opportunities and solutions," *IEEE Netw.*, vol. 35, no. 2, pp. 152–159, 2021.

[52] A. A. Khadir and S. A. H. Seno, "SDN-based offloading policy to reduce the delay in fog-vehicular networks," *Peer-to-Peer Netw. Appl.*, vol. 14, no. 3, pp. 1261–1275, 2021.

[53] X. Jiang, F. R. Yu, T. Song, and V. C. Leung, "Resource allocation of video streaming over vehicular networks: A survey, some research issues and challenges," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 5955–5975, 2022. DOI: 10.1109/TITS.2021.3065209.

[54] N. Ganeshkumar and S. Kumar, "OBU (On-Board Unit) wireless devices in VANET(s) for effective communication—A review," *Comput. Methods and Data Eng.*, pp. 191–202, 2021.

[55] H. R. Abdulshaheed, Z. T. Yaseen, A. M. Salman, and I. Al_Barazanchi, "A survey on the use of WiMAX and Wi-Fi on Vehicular Ad-hoc Networks (VANETs)," in *IOP Conf. Ser.: Mater. Sci. Eng.*, IOP Publishing, vol. 870, 2020, p. 012 122.

[56] G. Kostadinov and T. Atanasova, "Security policies for wireless and network infrastructure," in *Probl. Eng. Cybern. Robot.*, Bulgarian Academy of Sciences, 2020, pp. 14–19.

[57] V. Nampally and D. M. R. Sharma, "Information sharing standards in communication for VANET," *Int. J. Sci. Res. Comput. Sci. Appl. Manag. Stu.*, vol. 7, no. 4, pp. 2319–1953, 2018.

[58] M. Stepzinski and S. Sengupta, "Cybersecurity analysis in dedicated short-range communications in vehicular networks," in *Proc. 11th IEEE Ann. Ubiquitous Comput. Electron. Mobile Commun. Conf. UEMCON*, IEEE, 2020, pp. 0021–0027.

[59] N. Rajatheva, I. Atzeni, E. Bjornson, *et al.*, "White paper on broadband connectivity in 6G," *arXiv preprint arXiv:2004.14247*, 2020.

[60] L. Zhao, W. Zhao, A. Hawbani, *et al.*, "Novel online sequential learning-based adaptive routing for edge software-defined vehicular networks," *IEEE Trans. Wirel. Commun.*, vol. 20, no. 5, pp. 2991–3004, 2020.

[61] B. Ji, Y. Wang, K. Song, *et al.*, "A survey of computational intelligence for 6G: Key technologies, applications and trends," *IEEE Trans. Ind. Inform.*, 2021.

[62] S. Nayak and M. Narvekar, "Real-time vehicle navigation using modified A* algorithm," in *Proc. Int. Conf. Emerg. Trends Innov. ICT (ICEI)*, IEEE, 2017, pp. 116–122.

[63] G. Tobar, P. Galdames, C. Gutierrez-Soto, and P. Rodriguez-Moreno, "A batching location cloaking algorithm for location privacy protection," *Collaborative Technologies and Data Science in Smart City Applications*, pp. 26–36, 2018.

[64] C. Bettini, "Privacy protection in location-based services: A survey," in *Handbook of Mobile Data Privacy*, Springer, 2018, pp. 73–96.

[65] J. W. Kim, K. Edemacu, and B. Jang, "Privacy-preserving mechanisms for location privacy in mobile crowdsensing: A survey," *J. Netw. Comput. Appl.*, vol. 200, p. 103 315, 2022. DOI: https://doi.org/10.1016/j.jnca.2021.103315.

[66] A. R. Sani, M. U. Hassan, and J. Chen, "Privacy preserving machine learning for electric vehicles: A survey," *arXiv preprint arXiv:2205.08462*, 2022.

[67] X. Deng, T. Gao, N. Guo, C. Zhao, and J. Qi, "PCP: A pseudonym change scheme for location privacy preserving in VANETs," *Entropy*, vol. 24, no. 5, p. 648, 2022.

[68] Q. Huang, X. Xu, H. Chen, and L. Xie, "A vehicle trajectory privacy preservation method based on caching and dummy locations in the Internet of Vehicles," *Sensors*, vol. 22, no. 12, p. 4423, 2022.

[69] 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. Mob. Comput.*, vol. 21, no. 7, pp. 2436–2450, 2022.

[70] L. Wu, X. Wei, L. Meng, S. Zhao, and H. Wang, "Privacy-preserving location-based traffic density monitoring," *Conn Sci.*, vol. 34, no. 1, pp. 874–894, 2022.

[71] V. K. Yadav, N. Andola, S. Verma, and S. Venkatesan, "Anonymous and linkable location-based services," *IEEE Trans. Veh. Technol.*, vol. 71, no. 9, pp. 9397–9409, 2022.

[72] Q. Li, H. Wu, X. Wu, and L. Dong, "Multi-level location privacy protection based on differential privacy strategy in VANETs," in *2019 IEEE 89th Veh. Technol. Conf. (VTC2019-Spring)*, IEEE, 2019, pp. 1–5.

[73] S. Takagi, Y. Cao, Y. Asano, and M. Yoshikawa, "Geo-graph-indistinguishability: Protecting location privacy for lbs over road networks," in *IFIP Annual Conf. Data Appl. Secur. Priv.*, Springer, 2019, pp. 143–163.

[74] M. E. Andrés, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, "Geo-indistinguishability: Differential privacy for location-based systems," in *Proc. of the 2013 ACM SIGSAC Conf. Comput. commun. secur.*, 2013, pp. 901–914.

[75] I. Ullah, M. A. Shah, A. Khan, and G. Jeon, "Privacy-preserving multilevel obfuscation scheme for vehicular network," *Trans. Emerg. Telecommun. Technol.*, vol. 32, no. 2, e4204, 2021.

[76] T. Meuser, O. T. Ojo, D. Bischoff, A. F. Anta, I. Stavrakakis, and R. Steinmetz, "Hide me: Enabling location privacy in heterogeneous vehicular networks," in *Int. Conf. Netw. Syst.*, Springer, 2020, pp. 11–27.

[77] J. Hua, W. Tong, F. Xu, and S. Zhong, "A geo-indistinguishable location perturbation mechanism for location-based services supporting frequent queries," *IEEE Trans. Inf. Forens.. Secur.*, vol. 13, no. 5, pp. 1155–1168, 2017.

[78] Q. A. Arain, I. Memon, Z. Deng, M. H. Memon, F. A. Mangi, and A. Zubedi, "Location monitoring approach: Multiple mix-zones with location privacy protection based on traffic flow over road networks," *Multimed. Tools. Appl.*, vol. 77, no. 5, pp. 5563–5607, 2018.

[79] S. Sultanuddin and M. Ali Hussain, "Token system-based efficient route optimization in mobile ad hoc network for vehicular ad hoc network in smart city," *Trans. Emerg. Telecommun. Technol.*, vol. 31, no. 12, e3853, 2020.

[80] J. Kang, D. Steiert, D. Lin, and Y. Fu, "Movewithme: Location privacy preservation for smartphone users," *IEEE Trans. Inf. Forens. Secur.*, vol. 15, pp. 711–724, 2019.

[81] J. Zhang, F. Yang, Z. Ma, Z. Wang, X. Liu, and J. Ma, "A decentralized location privacy-preserving spatial crowdsourcing for internet of vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 4, pp. 2299–2313, 2020.

[82] L. Wang, D. Yang, X. Han, T. Wang, D. Zhang, and X. Ma, "Location privacy-preserving task allocation for mobile crowdsensing with differential geo-obfuscation," in *Proc. of the 26th Int. Conf. World Wide Web*, 2017, pp. 627–636.

[83] O. Abul and C. Bayrak, "From location to location pattern privacy in location-based services," *Knowl. Inf. Sys.*, vol. 56, no. 3, pp. 533–557, 2018.

[84] Y. Adegoke, "Uber drivers in lagos are using a fake gps app to inflate rider fares," *Quartz Africa, November*, vol. 13, 2017.

[85] Y. Ge, H. Xiong, C. Liu, and Z.-H. Zhou, "A taxi driving fraud detection system," in *2011 IEEE 11th International Conference on Data Mining*, IEEE, 2011, pp. 181–190.

[86] E. Ekenstedt, L. Ong, Y. Liu, S. Johnson, P. L. Yeoh, and J. Kliewer, "When differential privacy implies syntactic privacy," *IEEE Trans. Inf. Forens. Secur.*, vol. 17, pp. 2110–2124, 2022. DOI: 10.1109/TIFS.2022.3177953.

[87] B. Niu, Y. Chen, Z. Wang, F. Li, B. Wang, and H. Li, "Eclipse: Preserving differential location privacy against long-term observation attacks," *IEEE Trans. Mob. Comput.*, vol. 21, no. 1, pp. 125–138, 2020.

[88] L. Benarous, S. Bitam, and A. Mellouk, "Cslpps: Concerted silence-based location privacy preserving scheme for internet of vehicles," *IEEE Trans. Vel. Technol.*, vol. 70, no. 7, pp. 7153–7160, 2021.

[89] R. Al-Dhubhani and J. M. Cazalas, "An adaptive geo-indistinguishability mechanism for continuous lbs queries," *Wireless Networks*, vol. 24, no. 8, pp. 3221–3239, 2018.

[90] N. Kaaniche, M. Laurent, and S. Belguith, "Privacy enhancing technologies for solving the privacy-personalization paradox: Taxonomy and survey," *J. Netw. Comput. Appl.*, vol. 171, p. 102 807, 2020.

[91] S. Su, X. Zeng, S. Song, *et al.*, "Positioning accuracy improvement of automated guided vehicles based on a novel magnetic tracking approach," *IEEE Intell. Transp. Syst. Mag.*, vol. 12, no. 4, pp. 138–148, 2018.

[92] Y. Feng, G. Mao, B. Cheng, *et al.*, "MagMonitor: Vehicle speed estimation and vehicle classification through a magnetic sensor," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 2, pp. 1311–1322, 2022. DOI: 10.1109/TITS.2020.3024652.

[93] Z. Zhang, X. Mao, K. Zhou, and H. Yuan, "Collaborative sensing-based parking tracking system with wireless magnetic sensor network," *IEEE Sens. J.*, vol. 20, no. 9, pp. 4859–4867, 2020.

[94] Y. Feng, G. Mao, B. Cheng, B. Huang, S. Wang, and J. Chen, "MagSpeed: A novel method of vehicle speed estimation through a single magnetic sensor," in *Proc. IEEE Intell. Transp. Syst. Conf. (ITSC) 2019*, IEEE, 2019, pp. 4281–4286.

[95] S. Su, H. Dai, S. Cheng, P. Lin, C. Hu, and B. Lv, "A robust magnetic tracking approach based on graph optimization," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 10, pp. 7933–7940, 2020.

[96] C. Röger, M. Kalinic, and J. M. Krisp, "Extracting densely covered areas within floating car datasets using inductive loop detector data," *KN-J. Cartogr. Geogr. Inf.*, vol. 71, no. 2, pp. 97–103, 2021.

[97] R. A. Gheorghiu, V. Iordache, and V. A. Stan, "Urban traffic detectors-comparison between inductive loop and magnetic sensors," in *Proc. 13th Int. Conf. Electron. Comput. Artif. Intell. (ECAI) 2021)*, IEEE, 2021, pp. 1–4.

[98] C. Spiess, "Is that traffic light tracking you? A case study on a municipal surveillance technology in seattle," *IEEE Trans. Technol. Soc.*, vol. 2, no. 1, pp. 15–19, 2021.

[99] G. Nair, B. A. Kumar, and L. Vanajaskshi, "Mapping bus and stream travel time using machine learning approaches," *J. Adv. Transp.*, vol. 2022, 2022.

[100] A. Khadhir, H. Maripini, S. Sreedhar, L. Vanajakshi, and B. R. Chilukuri, "A study of delay estimation methods at signalized intersections for mixed traffic condition," *Transp. Dev. Econ.*, vol. 7, no. 1, pp. 1–16, 2021.

[101] Y. Alkendi, L. Seneviratne, and Y. Zweiri, "State of the art in vision-based localization techniques for autonomous navigation systems," *IEEE Access*, vol. 9, pp. 76 847–76 874, 2021.

[102] J. Feng, D. Zeng, X. Jia, *et al.*, "Cross-frame keypoint-based and spatial motion information-guided networks for moving vehicle detection and tracking in satellite videos," *ISPRS J. Photogramm. Remote Sens.*, vol. 177, pp. 116–130, 2021.

[103] H. Yang, J. Cai, M. Zhu, C. Liu, and Y. Wang, "Traffic-Informed Multi-camera Sensing (TIMS) system based on vehicle re-identification," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 10, pp. 17 189–17 200, 2022. DOI: 10.1109/TITS.2022.3154368.
[104] S. Pooja, "Vehicle tracking system using GPS," *Int. J. Sci. Res.*, vol. 2, no. 9, pp. 128–130, 2013.
[105] E. Levy, A. Shabtai, B. Groza, P.-S. Murvay, and Y. Elovici, "CAN-LOC: Spoofing detection and physical intrusion localization on an in-vehicle CAN bus based on deep features of voltage signals," *arXiv preprint arXiv:2106.07895*, 2021.
[106] A. Eldesoky, A. M. Kamel, M. Elhabiby, and H. Elhennawy, "Real time localization solution for land vehicle application using low-cost integrated sensors with GPS," *J. Appl. Res. Technol.*, vol. 18, no. 4, pp. 214–228, 2020.
[107] C. R. Casal, "Privacy within in-car systems," *Info.*, vol. 7, no. 6, pp. 66–75, 2005.
[108] N. Moayeri, J. Mapar, S. Tompkins, and K. Pahlavan, "Emerging opportunities for localization and tracking [Guest Editorial]," *IEEE Wirel. Commun.*, vol. 18, no. 2, pp. 8–9, 2011. DOI: 10.1109/MWC.2011.5751290.
[109] I. Rouf, R. Miller, H. Mustafa, *et al.*, "Security and privacy vulnerabilities of in-car wireless networks: A tire pressure monitoring system case study," in *Proc. 19th USENIX Secur. Symp. (USENIX Security 10)*, 2010.
[110] Y. Watanabe, H. Yamamoto, and H. Yoshida, "A study on the applicability of the Lesamnta-LW lightweight hash function to TPMS," *IACR Cryptol. ePrint Arch.*, vol. 2019, p. 840, 2019.
[111] M. Page and T. L. Wickramarathne, "Enhanced situational awareness with signals of opportunity: RSS-based localization and tracking," in *Proc. IEEE Intell. Transp. Syst. Conf. (ITSC) 2019*, IEEE, 2019, pp. 3833–3838.
[112] J. Han, E. Owusu, L. T. Nguyen, A. Perrig, and J. Zhang, "ACComplice: Location inference using accelerometers on smartphones," in *Proc. 4th Int. Conf. Commun. Syst. Networks (COMSNETS) 2012*, IEEE, 2012, pp. 1–9.
[113] Z. Li, Q. Pei, I. Markwood, Y. Liu, M. Pan, and H. Li, "Location privacy violation via GPS-agnostic smart phone car tracking," *IEEE Trans. Veh. Technol.*, vol. 67, no. 6, pp. 5042–5053, 2018.
[114] S. Guha, K. Plarre, D. Lissner, *et al.*, "AutoWitness: Locating and tracking stolen property while tolerating GPS and radio outages," *ACM Trans. Sens. Netw.*, vol. 8, no. 4, pp. 1–28, 2012.
[115] H. Guo, X. Zhou, J. Liu, and Y. Zhang, "Vehicular intelligence in 6G: Networking, communications, and computing," *Veh. Commun.*, vol. 33, p. 100 399, 2022.
[116] S. V. Pushpakaran, J. M. Purushothama, M. Mani, A. Chandroth, M. Pezholil, and V. Kesavath, "A metamaterial absorber based high gain directional dipole antenna," *Int. J. Microw. Wirel. Technol.*, vol. 10, no. 4, pp. 430–436, 2018.
[117] D. Burghal, G. Phadke, A. Nair, *et al.*, "Supervised learning approach for relative vehicle localization using V2V MIMO links," in *Proc. IEEE Int. Conf. Commun. (ICC) 2022*, IEEE, 2022, pp. 4528–4534.
[118] M. K. Emara, D. J. King, H. V. Nguyen, S. Abielmona, and S. Gupta, "Millimeter-Wave slot array antenna front-end for amplitude-only direction finding," *IEEE Trans. Antennas Propag.*, vol. 68, no. 7, pp. 5365–5374, 2020.
[119] K. Heimann, J. Tiemann, S. Böcker, and C. Wietfeld, "Cross-bearing based positioning as a feature of 5G millimeter wave beam alignment," in *Proc. IEEE 91st Vehic. Technol. Conf. (VTC2020-Spring)*, IEEE, 2020, pp. 1–5.

[120] P. Bagga, A. K. Das, M. Wazid, J. J. Rodrigues, and Y. Park, "Authentication protocols in Internet of Vehicles: Taxonomy, analysis, and challenges," *IEEE Access*, vol. 8, pp. 54 314–54 344, 2020.
[121] I. Saini, B. St Amour, and A. Jaekel, "Intelligent adversary placements for privacy evaluation in VANET," *Information*, vol. 11, no. 9, p. 443, 2020.
[122] R. Hussain, D. Kim, J. Son, *et al.*, "Secure and privacy-aware incentives-based witness service in social Internet of Vehicles clouds," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2441–2448, 2018.
[123] A. Boualouache, S.-M. Senouci, and S. Moussaoui, "Towards an efficient pseudonym management and changing scheme for vehicular ad-hoc networks," in *IEEE Glob. Commun. Conf. (GLOBECOM) 2016*, IEEE, 2016, pp. 1–7.
[124] A. Boualouache and S. Moussaoui, "TAPCS: Traffic-aware pseudonym changing strategy for VANETs," *Peer-to-Peer Netw. Appl.*, vol. 10, no. 4, pp. 1008–1020, 2017.
[125] S. Wang and N. Yao, "A RSU-aided distributed trust framework for pseudonym-enabled privacy preservation in VANETs," *Wirel. Netw.*, vol. 25, no. 3, pp. 1099–1115, 2019.
[126] A. Boualouache and S. Moussaoui, "Urban pseudonym changing strategy for location privacy in VANETs," *Int. J. Ad Hoc Ubiquitous Comput.*, vol. 24, no. 1-2, pp. 49–64, 2017.
[127] D. Das, S. Banerjee, U. Ghosh, U. Biswas, and A. K. Bashir, "A decentralized vehicle anti-theft system using blockchain and smart contracts," *Peer-to-Peer Netw. Appl.*, vol. 14, no. 5, pp. 2775–2788, 2021.
[128] J. Zhang, W. Xiao, B. Coifman, and J. P. Mills, "Vehicle tracking and speed estimation from roadside lidar," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 5597–5608, 2020.
[129] S. D. Khan and H. Ullah, "A survey of advances in vision-based vehicle re-identification," *Comput. Vis. Image Underst.*, vol. 182, pp. 50–63, 2019.
[130] P. Khoenkaw, "An implementation of programmable oscillator for inductive-loop vehicle sensor using low-cost microcontroller," in *2021 Joint Int. Conf. Digital Arts Media Technol. ECTI North. Sect. Conf. Electr. Electron. Comput. Telecommun. Eng.*, IEEE, 2021, pp. 31–36.
[131] M. W. Raad, M. Deriche, and T. Sheltami, "An IoT-based school bus and vehicle tracking system using RFID technology and mobile data networks," *Arab. J. Sci. Eng.*, vol. 46, no. 4, pp. 3087–3097, 2021.
[132] N. Ismail, S. Wahid, and N. Ahmad, "Arduino based RFID vehicle tracking for home security," in *J. Phys.: Conf. Ser.*, IOP Publishing, vol. 1529, 2020, p. 022 060.
[133] A. Motroni, A. Buffi, P. Nepa, and B. Tellini, "Sensor-fusion and tracking method for indoor vehicles with low-density UHF-RFID tags," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–14, 2020.
[134] Y. Zhang, X. Gong, K. Liu, and S. Zhang, "Localization and tracking of an indoor autonomous vehicle based on the phase difference of passive UHF RFID signals," *Sensors*, vol. 21, no. 9, p. 3286, 2021.
[135] W.-C. Ma, I. Tartavull, I. A. Bârsan, *et al.*, "Exploiting sparse semantic HD maps for self-driving vehicle localization," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst. IROS 2019*, IEEE, 2019, pp. 5304–5311.
[136] Z. Xiao, P. Li, V. Havyarimana, G. M. Hassana, D. Wang, and K. Li, "GOI: A novel design for vehicle positioning and trajectory prediction under urban environments," *IEEE Sens. J.*, vol. 18, no. 13, pp. 5586–5594, 2018.

[137] X. Lin, B. Ma, X. Wang, Y. He, R. P. Liu, and W. Ni, "Multi-layer reverse engineering system for vehicular controller area network messages," in *Proc. 2022 IEEE 25th Int. Conf. Comput. Support. Coop. Work Des.*, IEEE, 2022, pp. 1185–1190.

[138] S.-F. Lokman, A. T. Othman, and M.-H. Abu-Bakar, "Intrusion detection system for automotive controller area network (CAN) bus system: A review," *EURASIP. J. Wirel. Commun. Netw.*, vol. 2019, pp. 1–17, 2019.

[139] A. H. Kelechi, M. H. Alsharif, D. A. Oluwole, *et al.*, "The recent advancement in unmanned aerial vehicle tracking antenna: A review," *Sensors*, vol. 21, no. 16, p. 5662, 2021.

[140] M. A. G. Al-Sadoon, R. Asif, Y. I. A. Al-Yasir, R. A. Abd-Alhameed, and P. S. Excell, "AoA localization for vehicle-tracking systems using a dual-band sensor array," *IEEE Trans. Antennas Propag.*, vol. 68, no. 8, pp. 6330–6345, 2020.

[141] J. Jiang, G. Han, H. Wang, and M. Guizani, "A survey on location privacy protection in wireless sensor networks," *J. Netw. Comput. Appl.*, vol. 125, pp. 93–114, 2019.

[142] M. Bradbury, P. Taylor, U. I. Atmaca, C. Maple, and N. Griffiths, "Privacy challenges with protecting live vehicular location context," *IEEE Access*, vol. 8, pp. 207 465–207 484, 2020.

[143] K. Emara, W. Woerndl, and J. Schlichter, "On evaluation of location privacy preserving schemes for VANET safety applications," *Comput. Commun.*, vol. 63, pp. 11–23, 2015.

[144] A. Pandey and S. Yadav, "Performance evaluation of amplify-and-forward relaying cooperative vehicular networks under physical layer security," *Trans. Emerg. Telecommun. Technol.*, vol. 29, no. 12, e3534, 2018.

[145] S. Shen, D. Chen, Y.-L. Wei, Z. Yang, and R. R. Choudhury, "Voice localization using nearby wall reflections," in *Proc. 26th Ann. Int. Conf. Mobile Comput. Netw.*, 2020, pp. 1–14.

[146] A. Bahramali, R. Soltani, A. Houmansadr, D. Goeckel, and D. Towsley, "Practical traffic analysis attacks on secure messaging applications," *arXiv preprint arXiv:2005.00508*, 2020.

[147] N. Lal, S. Kumar, and V. K. Chaurasiya, "A road monitoring approach with real-time capturing of events for efficient vehicles safety in smart city," *Wirel. Pers. Commun.*, vol. 114, no. 1, pp. 657–674, 2020.

[148] S. Zakhary and A. Benslimane, "On location-privacy in opportunistic mobile networks, a survey," *J. Netw. Comput. Appl.*, vol. 103, pp. 157–170, 2018.

[149] J. Zhou, Z. Cao, X. Dong, and A. V. Vasilakos, "Security and privacy for cloud-based IoT: Challenges," *IEEE Commun. Mag.*, vol. 55, no. 1, pp. 26–33, 2017.

[150] M. Chaurasia and B. P. Singh, "Prevention of DoS and routing attack in OLSR under MANET," in *Proc. Int. Conf. Recent Advancement Comp. Commun.*, Springer, 2018, pp. 287–295.

[151] L. Zhu, S. Wang, C. Li, and Z. Yang, "License plate recognition in urban road based on vehicle tracking and result integration," *J. Intell. Syst.*, vol. 29, no. 1, pp. 1587–1597, 2020.

[152] Y. Qian, L. Yu, W. Liu, and A. G. Hauptmann, "Electricity: An efficient multi-camera vehicle tracking system for intelligent city," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, 2020, pp. 588–589.

[153] J. Liu, P. Jayakumar, J. L. Stein, and T. Ersal, "Combined speed and steering control in high-speed autonomous ground vehicles for obstacle avoidance using model predictive control," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 8746–8763, 2017.

[154] S. Nie, L. Liu, and Y. Du, "Free-fall: Hacking Tesla from wireless to CAN bus," in *Briefing, Black Hat USA*, vol. 25, 2017, pp. 1–16.
[155] M. H. Eiza and Q. Ni, "Driving with sharks: Rethinking connected vehicles with vehicle cybersecurity," *IEEE Veh. Technol. Mag.*, vol. 12, no. 2, pp. 45–51, 2017.
[156] R. Molina-Masegosa, M. Sepulcre, J. Gozalvez, F. Berens, and V. Martinez, "Empirical models for the realistic generation of cooperative awareness messages in vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5713–5717, 2020.
[157] E. Ahmed and H. Gharavi, "Cooperative vehicular networking: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 3, pp. 996–1014, 2018.
[158] B. Lonc, F. Haidar, and D. Filatov, "Cooperative ITS security standards: Implementation, assessment and next challenges," in *ITS European Congress*, 2020.
[159] N. Padmapriya and S. Subathra, "A secure communication and location privacy in WSN using grey wolf optimization," in *J. Phys.: Conf. Ser.*, IOP Publishing, vol. 1717, 2021, p. 012 050.
[160] S. Alharthi, P. Johnson, and M. Randles, "Secure and energy-efficient communication in IoT/CPS," *Recent Trends in Commun. Netw.*, 2020.
[161] V. Sadhu, S. Zonouz, V. Sritapan, and D. Pompili, "CollabLoc: Privacy-preserving multimodal collaborative mobile phone localization," *IEEE Trans. Mob. Comput.*, vol. 20, no. 1, pp. 104–116, 2019.
[162] B. Hong, S. Bae, and Y. Kim, "GUTI reallocation demystified: Cellular location tracking with changing temporary identifier.," in *NDSS*, 2018.
[163] R. Lu, L. Zhang, J. Ni, and Y. Fang, "5G vehicle-to-everything services: Gearing up for security and privacy," *Proc. IEEE*, vol. 108, no. 2, pp. 373–389, 2019.
[164] Q. Zhao, H. Wen, Z. Lin, D. Xuan, and N. Shroff, "On the accuracy of measured proximity of bluetooth-based contact tracing apps," in *Proc. Int. Conf. Security Priv. Commun. Syst*, Springer, 2020, pp. 49–60.
[165] Y. Xu, D. He, P. Chao, J. Kim, W. Hua, and X. Zhou, "Route reconstruction using low-quality bluetooth readings," in *Proc. 28th Int. Conf. Adv. Geogr. Inf. Syst.*, 2020, pp. 179–182.
[166] P. Verde, J. Dıéez-González, R. Ferrero-Guillén, A. Martıénez-Gutiérrez, and H. Perez, "Memetic chains for improving the local wireless sensor networks localization in urban scenarios," *Sensors*, vol. 21, no. 7, p. 2458, 2021.
[167] S. Y. Fayi and Z. Sheng, "A survey of security, privacy and trust issues in vehicular computation offloading and their solutions using blockchain.," *Open Research Europe*, vol. 3, 2023.
[168] A. Haddaji, S. Ayed, and L. C. Fourati, "Artificial intelligence techniques to mitigate cyber-attacks within vehicular networks: Survey," *Computers and Electrical Engineering*, vol. 104, p. 108 460, 2022.
[169] T. Alladi, V. Chamola, N. Sahu, V. Venkatesh, A. Goyal, and M. Guizani, "A comprehensive survey on the applications of blockchain for securing vehicular networks," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 1212–1239, 2022.
[170] M. Platt and P. McBurney, "Sybil in the haystack: A comprehensive review of blockchain consensus mechanisms in search of strong sybil attack resistance," *Algorithms*, vol. 16, no. 1, p. 34, 2023.
[171] A. Altaweel, H. Mukkath, and I. Kamel, "Gps spoofing attacks in fanets: A systematic literature review," *IEEE Access*, 2023.

[172] S. Jena, N. P. Padhy, and J. M. Guerrero, "Multi-layered coordinated countermeasures for dc microgrid clusters under man in the middle attack," *IEEE Transactions on Industry Applications*, 2023.

[173] N. Sidhu and M. Sachdeva, "A comprehensive study of routing layer intrusions in zigbee based wireless sensor networks," *International Journal of Advanced Science and Technology*, vol. 29, no. 3, pp. 514–524, 2020.

[174] E. M. Mianji, G.-M. Muntean, and I. Tal, "Trustworthy routing in vanet: A q-learning approach to protect against black hole and gray hole attacks," in *2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*, IEEE, 2023, pp. 1–6.

[175] R. R. Khan, A. Hanif, and Q. Ahmed, "Threat analysis of position, navigation, and timing for highly automated vehicles," in *2023 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, IEEE, 2023, pp. 647–659.

[176] A. Rasheed, M. Baza, M. Badr, H. Alshahrani, and K.-K. R. Choo, "Efficient crypto engine for authenticated encryption, data traceability, and replay attack detection over can bus network," *IEEE Transactions on Network Science and Engineering*, 2023.

[177] R. Prathap Kumar, U. Srilakshmi, and K. Ganesh Reddy, "Routing integrity mechanism to prevent wormhole attacks in vehicular adhoc networks," in *Inventive Systems and Control: Proceedings of ICISC 2023*, Springer, 2023, pp. 53–64.

[178] A. Mairaj and A. Y. Javaid, "Game theoretic solution for an unmanned aerial vehicle network host under ddos attack," *Computer Networks*, vol. 211, p. 108 962, 2022.

[179] C. Dunbar and G. Qu, "A DTN routing protocol for vehicle location information protection," in *Proc. - IEEE Mil. Commun. Conf. 2014*, IEEE, 2014, pp. 94–100.

[180] C. Dunbar, M. Gao, and G. Qu, "Pass and run: A privacy preserving delay tolerant network communication protocol for cybervehicles," in *2013 Int. Conf. Connect. Veh. Expo. (ICCVE)*, IEEE, 2013, pp. 840–841.

[181] Z. Lu, M. Gao, Z. Liu, G. Qu, and C. Dunbar, "Pass and run: A privacy preserving delay tolerant network communication protocol for cybervehicles," *IEEE Des. Test*, vol. 36, no. 6, pp. 56–62, 2019.

[182] Y. Qin, J. Hao, G. Han, J. Pan, K. Han, and H. Niu, "Research on secured communication of intelligent connected vehicle based on digital certificate," in *Intell. Autom. Soft Comput.*, Springer, 2021, pp. 1068–1080.

[183] S. Terzi, C. Savvaidis, K. Votis, D. Tzovaras, and I. Stamelos, "Securing emission data of smart vehicles with blockchain and self-sovereign identities," in *Proc. IEEE Int. Conf. Blockchain 2020*, IEEE, 2020, pp. 462–469.

[184] M. Langheinrich, "Privacy by design—principles of privacy-aware ubiquitous systems," in *Proc. Int. Conf. Ubiquitous Comput.*, Springer, 2001, pp. 273–291.

[185] S. F. Shahandashti and R. Safavi-Naini, "Threshold attribute-based signatures and their application to anonymous credential systems," in *Proc. AFRICACRYPT 2009, 2nd Int. Conf. Cryptol. Africa*, Springer, 2009, pp. 198–216.

[186] H. K. Maji, M. Prabhakaran, and M. Rosulek, "Attribute-based signatures," in *Proc. Cryptograph. Track RSA Conf. (CT-RSA)*, Springer, 2011, pp. 376–392.

[187] T. Okamoto and K. Takashima, "Efficient attribute-based signatures for non-monotone predicates in the standard model," *IEEE Trans. Cloud Comput.*, vol. 2, no. 4, pp. 409–421, 2014.

[188] J. Li, M. H. Au, W. Susilo, D. Xie, and K. Ren, "Attribute-based signature and its applications," in *Proc. 5th ACM Symp. Info. Comput. Commun. Secur.*, 2010, pp. 60–69.

[189] Y. Zhang and D. Feng, "Efficient attribute proofs in anonymous credential using attribute-based cryptography," in *Int. Conf. Inf. Commun. Secur. (ICICS)*, Springer, 2012, pp. 408–415.

[190] A. E. Kaafarani, E. Ghadafi, and D. Khader, "Decentralized traceable attribute-based signatures," in *Proc. Cryptograph. Track RSA Conf. (CT-RSA)*, Springer, 2014, pp. 327–348.

[191] N. Kaaniche and M. Laurent, "Attribute-based signatures for supporting anonymous certification," in *Proc. Eur. Symp. Res. Comput. Secur. (ESORICS)*, Springer, 2016, pp. 279–300.

[192] R. E. Bansarkhani and R. Misoczki, "G-Merkle: A hash-based group signature scheme from standard assumptions," in *Proc. Int. Conf. Post-Quantum Cryptogr.*, Springer, 2018, pp. 441–463.

[193] H. Zheng, Q. Wu, B. Qin, L. Zhong, S. He, and J. Liu, "Linkable group signature for auditing anonymous communication," in *Proc. Australasian Conf. Inf. Secur. Privacy (ACISP)*, Springer, 2018, pp. 304–321.

[194] S. A. A. Hakeem and H. Kim, "Multi-zone Authentication and Privacy-preserving Protocol (MAPP) based on the bilinear pairing cryptography for 5G-V2X," *Sensors*, vol. 21, no. 2, p. 665, 2021.

[195] L. Wu, J. Fan, Y. Xie, J. Wang, and Q. Liu, "Efficient location-based conditional privacy-preserving authentication scheme for vehicle ad hoc networks," *Int. J. Distrib. Sens. Netw.*, vol. 13, no. 3, p. 1 550 147 717 700 899, 2017.

[196] P. Mundhe, V. K. Yadav, A. Singh, S. Verma, and S. Venkatesan, "Ring signature-based conditional privacy-preserving authentication in VANETs," *Wirel. Pers. Commun.*, vol. 114, no. 1, pp. 853–881, 2020.

[197] P. Mundhe, V. K. Yadav, S. Verma, and S. Venkatesan, "Efficient lattice-based ring signature for message authentication in VANETs," *IEEE Syst. J.*, vol. 14, no. 4, pp. 5463–5474, 2020.

[198] Q. Mei, H. Xiong, J. Chen, M. Yang, S. Kumari, and M. K. Khan, "Efficient certificateless aggregate signature with conditional privacy preservation in IoV," *IEEE Syst. J.*, vol. 15, no. 1, pp. 245–256, 2020.

[199] G. Ateniese, D. H. Chou, B. d. Medeiros, and G. Tsudik, "Sanitizable signatures," in *Proc. Eur. Symp. Res. Comput. Secur. (ESORICS)*, Springer, 2005, pp. 159–177.

[200] D. Pàmies-Estrems, N. Kaaniche, M. Laurent, J. Castellà-Roca, and J. Garcia-Alfaro, "Lifelogging protection scheme for internet-based personal assistants," in *Data Privacy Management, Cryptocurrencies and Blockchain Technology*, Springer, 2018, pp. 431–440.

[201] H. Wang, J. Gan, Y. Feng, Y. Li, and X. Fu, "A privacy enhancement scheme based on blockchain and blind signature for Internet of Vehicles," in *Proc. Int. Conf. Blockchain Trustw. Syst.*, Springer, 2021, pp. 368–387.

[202] G. Sun, S. Sun, H. Yu, and M. Guizani, "Toward incentivizing fog-based privacy-preserving mobile crowdsensing in the Internet of Vehicles," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 4128–4142, 2019.

[203] A. C. Yao, "Protocols for secure computations," in *Proc. 23rd Ann. Symp. Found. Comp. Sci. (SFCS 1982)*, IEEE, 1982, pp. 160–164.

[204] J. W. Kim, K. Edemacu, J. S. Kim, Y. D. Chung, and B. Jang, "A survey of differential privacy-based techniques and their applicability to location-based services," *Comput. Secur.*, vol. 111, p. 102 464, 2021.

[205] B. Niu, Q. Li, X. Zhu, G. Cao, and H. Li, "Achieving k-anonymity in privacy-aware location-based services," in *IEEE INFOCOM 2014 - IEEE Conf. Comput. Commun.*, IEEE, 2014, pp. 754–762.

[206] H. Liu, X. Li, H. Li, J. Ma, and X. Ma, "Spatiotemporal correlation-aware dummy-based privacy protection scheme for location-based services," in *IEEE INFOCOM 2017 - IEEE Conf. Comput. Commun.*, 2017, pp. 1–9.

[207] Y. Xiao and L. Xiong, "Protecting locations with differential privacy under temporal correlations," in *Proc. 22nd ACM SIGSAC Conf. Comput. Commun. Secur.*, 2015, pp. 1298–1309.

[208] M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking," in *Proc. 1st Int. Conf. Mobile Syst. Appl. Serv.*, 2003, pp. 31–42.

[209] J. C. Duchi, M. I. Jordan, and M. J. Wainwright, "Local privacy and statistical minimax rates," in *Proc. 2013 IEEE 54th Ann. Symp. Foundations Comput. Sci. (FOCS)*, IEEE, 2013, pp. 429–438.

[210] Ú. Erlingsson, V. Feldman, I. Mironov, A. Raghunathan, K. Talwar, and A. Thakurta, "Amplification by shuffling: From local to central differential privacy via anonymity," in *Proc. 30th Ann. ACM-SIAM Symp. Discrete Algorithms*, SIAM, 2019, pp. 2468–2479.

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

[212] 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, 2018.

[213] Q. Li, H. Wu, X. Wu, and L. Dong, "Multi-level location privacy protection based on differential privacy strategy in VANETs," in *Proc. IEEE 89th Veh. Technol. Conf. (VTC2019-Spring)*, 2019, pp. 1–5.

[214] X. Li, Y. Ren, L. T. Yang, *et al.*, "Perturbation-hidden: Enhancement of vehicular privacy for location-based services in Internet of Vehicles," *IEEE Trans. Netw. Sci. Eng.*, pp. 1–1, 2020.

[215] S. Haider, D. Gao, R. Ali, A. Hussain, and M. T. Ikram, "A privacy conserves pseudonym acquisition scheme in vehicular communication systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 9, pp. 15 536–15 545, 2022.

[216] M. A. Al-Shareeda, M. Anbar, S. Manickam, and I. H. Hasbullah, "A secure pseudonym-based conditional privacy-preservation authentication scheme in vehicular ad hoc networks," *Sensors*, vol. 22, no. 5, p. 1696, 2022.

[217] S. Wang, N. Yao, N. Gong, and Z. Gao, "A trigger-based pseudonym exchange scheme for location privacy preserving in VANETs," *Peer-to-Peer Netw. Appl.*, vol. 11, no. 3, pp. 548–560, 2018.

[218] X. Li, H. Zhang, Y. Ren, *et al.*, "PAPU: Pseudonym swap with provable unlinkability based on differential privacy in VANETs," *IEEE Internet Things J.*, vol. 7, no. 12, pp. 11 789–11 802, 2020. DOI: 10.1109/JIOT.2020.3001381.

[219] X. Deng, T. Gao, N. Guo, J. Qi, and C. Zhao, "PAS: Privacy-preserving authentication scheme based on SDN for VANETs," *Appl. Sci.*, vol. 12, no. 9, p. 4791, 2022.

[220] P. Ma, D. Tao, and T. Wu, "A pseudonym based anonymous identity authentication mechanism for mobile crowd sensing," in *Proc. 3rd Int. Conf. Big Data Comput. Commun. (BIGCOM)*, IEEE, 2017, pp. 10–14.

[221] H. Li, X. Xue, Z. Li, L. Li, and J. Xiong, "Location privacy protection scheme for LBS in IoT," *Wirel. Commun. Mob. Comput.*, vol. 2021, pp. 1–18, 2021.

[222] 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. (CSUR)*, vol. 54, no. 1, pp. 1–36, 2021.

[223] T. Carvalho, N. Moniz, P. Faria, and L. Antunes, "Survey on privacy-preserving techniques for data publishing," *arXiv preprint arXiv:2201.08120*, 2022.

[224] S. Zhang, X. Li, Z. Tan, T. Peng, and G. Wang, "A caching and spatial k-anonymity driven privacy enhancement scheme in continuous location-based services," *Future Gener. Comput. Syst.*, vol. 94, pp. 40–50, 2019.

[225] P.-s. Xie, X.-m. Han, T. Feng, Y. Yan, and G.-q. Ma, "Location privacy protection algorithm based on the division method of voronoi in the Internet of Vehicles," in *Proc. Int. Conf. Artif. Intell. Secur.*, Springer, 2020, pp. 412–422.

[226] M. Afifi, K. Zhou, and J. Ren, "Privacy characterization and quantification in data publishing," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 9, pp. 1756–1769, 2018.

[227] B. Sowmiya and E. Poovammal, "A heuristic k-anonymity based privacy preserving for student management hyperledger fabric blockchain," *Wirel. Pers. Commun.*, vol. 2021, pp. 1–18, 2021.

[228] K. Arava and S. Lingamgunta, "Adaptive k-anonymity approach for privacy preserving in cloud," *Arab. J. Sci. Eng.*, vol. 45, no. 4, pp. 2425–2432, 2020.

[229] H. Li, L. Pei, D. Liao, G. Sun, and D. Xu, "Blockchain meets VANET: An architecture for identity and location privacy protection in VANET," *Peer-to-Peer Netw. Appl.*, vol. 12, no. 5, pp. 1178–1193, 2019.

[230] S. Sharma and B. Kaushik, "Applications and challenges in Internet of Vehicles: A survey," *Internet Things its Appl.*, pp. 55–65, 2022.

[231] A. Boualouache, S.-M. Senouci, and S. Moussaoui, "A survey on pseudonym changing strategies for vehicular ad-hoc networks," *IEEE Commun. Surv. Tutor.*, vol. 20, no. 1, pp. 770–790, 2017.

[232] C. Wang, P. Zhao, H. Huang, R. Zhang, and W. Zhu, "Privacy-preserving techniques for the 5G-enabled location-based services," in *5G-Enabled Internet of Things*, CRC Press, 2019, pp. 269–299.

[233] Z. Zhao, A. Ye, L. Meng, and Q. Zhang, "Pseudonym changing for vehicles in VANETs: A game-theoretic analysis based approach," in *Proc. Int. Conf. Networking Network Appl. (NaNA) 2019*, IEEE, 2019, pp. 70–74.

[234] I. Memon, Q. Ali, A. Zubedi, and F. A. Mangi, "DPMM: Dynamic pseudonym-based multiple mix-zones generation for mobile traveler," *Multimed. Tools Appl.*, vol. 76, no. 22, pp. 24 359–24 388, 2017.

[235] B. Amro, "Protecting privacy in VANETs using mix zones with virtual pseudonym change," *arXiv preprint arXiv:1801.10294*, 2018.

[236] Q. Li, H. Wu, L. Liu, B. Pan, and L. Dong, "A group based dynamic mix zone scheme for location privacy preservation in VANETs," in *Proc. 3rd Int. Conf. Secur. Smart Cities, Ind. Control Syst. Commun. (SSIC) 2018*, 2018, pp. 1–5.

[237] J. Meng, X. Zhang, T. Cao, and Y. Xie, "Lightweight and anonymous mutual authentication protocol for IoT devices with physical unclonable functions," *Secur. Commun. Netw.*, vol. 2022, pp. 1–11, 2021.

[238] 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, 2020.

[239] P. Singh, A. Agarwal, G. Nakum, D. Rawat, and S. Nandi, "MPFSLP: Masqueraded probabilistic flooding for source-location privacy in VANETs," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11 383–11 393, 2020.

[240] S. Ghane, A. Jolfaei, L. Kulik, K. Ramamohanarao, and D. Puthal, "Preserving privacy in the Internet of Connected Vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 8, pp. 5018–5027, 2020.

[241] Q. Miao, W. Jing, and H. Song, "Differential privacy–based location privacy enhancing in edge computing," *Concurr. Comput. Pract. Exp.*, vol. 31, no. 8, e4735, 2019.

[242] S. Han, U. Topcu, and G. J. Pappas, "An approximately truthful mechanism for electric vehicle charging via joint differential privacy," in *Proc. Am. Control Conf. (ACC) 2015*, IEEE, 2015, pp. 2469–2475.

[243] W. Tong, J. Hua, and S. Zhong, "A jointly differentially private scheduling protocol for ridesharing services," *IEEE Trans. Inf. Forensics Secur.*, vol. 12, no. 10, pp. 2444–2456, 2017.

[244] X. Sun, F. R. Yu, P. Zhang, W. Xie, and X. Peng, "A survey on secure computation based on homomorphic encryption in vehicular ad hoc networks," *Sensors*, vol. 20, no. 15, p. 4253, 2020.

[245] A. Wood, K. Najarian, and D. Kahrobaei, "Homomorphic encryption for machine learning in medicine and bioinformatics," *ACM Comput. Surv. (CSUR)*, vol. 53, no. 4, pp. 1–35, 2020.

[246] Z. H. Mahmood and M. K. Ibrahem, "New fully homomorphic encryption scheme based on multistage partial homomorphic encryption applied in cloud computing," in *Proc. 1st Ann. Int. Conf. AiCIS*, IEEE, 2018, pp. 182–186.

[247] F. Yucel, K. Akkaya, and E. Bulut, "Efficient and privacy preserving supplier matching for electric vehicle charging," *Ad Hoc Netw.*, vol. 90, p. 101 730, 2019. DOI: https://doi.org/10.1016/j.adhoc.2018.07.029.

[248] V. Subramaniyaswamy, V. Jagadeeswari, V. Indragandhi, *et al.*, "Somewhat homomorphic encryption: Ring learning with error algorithm for faster encryption of IoT sensor signal-based edge devices," *Secur. Commun. Netw.*, vol. 2022, 2022, ISSN: 1939-0114. DOI: 10.1155/2022/2793998.

[249] H. Yu, H. Zhang, X. Yu, X. Du, and M. Guizani, "PGRide: Privacy-preserving group ridesharing matching in online ride hailing services," *IEEE Internet Things J.*, vol. 8, no. 7, pp. 5722–5735, 2020.

[250] S. Gupta and G. Arora, "Use of homomorphic encryption with GPS in location privacy," in *Proc. 4th Int. Conf. ISCON*, IEEE, 2019, pp. 42–45.

[251] N. Prema, "Efficient secure aggregation in VANETs using Fully Homomorphic Encryption (FHE)," *Mob. Netw. Appl.*, vol. 24, no. 2, pp. 434–442, 2019.

[252] M. A. Mohammed, B. Garcia-Zapirain, J. Nedoma, R. Martinek, P. Tiwari, N. Kumar, *et al.*, "Fully homomorphic enabled secure task offloading and scheduling system for transport applications," *IEEE Trans. Veh. Technol.*, vol. 71, no. 11, pp. 12 140–12 153, 2022.

[253] B. Chor, O. Goldreich, E. Kushilevitz, and M. Sudan, "Private information retrieval," in *Proc. 36th Annu. IEEE Found. Comput. Sci.*, IEEE, 1995, pp. 41–50.

[254] W. Gasarch, "A survey on private information retrieval," *Bulletin of the EATCS*, vol. 82, no. 72-107, p. 113, 2004.

[255] H. Sun and S. A. Jafar, "The capacity of private information retrieval," *IEEE Trans. Inf. Theory.*, vol. 63, no. 7, pp. 4075–4088, 2017.

[256] Z. Tan, C. Wang, M. Zhou, and L. Zhang, "Private information retrieval in vehicular location-based services," in *Proc. 4th IEEE World Forum on IoT (WF-IoT)*, IEEE, 2018, pp. 56–61.
[257] D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in *Proc. 2000 IEEE Symp. Secur. Priv. S&P 2000*, IEEE, 2000, pp. 44–55.
[258] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: Improved definitions and efficient constructions," *J. Comput. Secur.*, vol. 19, no. 5, pp. 895–934, 2011.
[259] M. A. Ferrag and A. Ahmim, "ESSPR: An efficient secure routing scheme based on searchable encryption with vehicle proxy re-encryption for vehicular peer-to-peer social network," *Telecommun. Syst.*, vol. 66, no. 3, pp. 481–503, 2017.
[260] Q. Tang, J. Shen, Z. Cao, and X. Dong, "PSSBP: A privacy-preserving scope-query searchable encryption scheme based on blockchain for parking lots sharing in vehicular networks," in *Proc. 19th IEEE Int. Conf. Embedded Ubiquitous Comput. (EUC)*, IEEE, 2021, pp. 1–8.
[261] Y. Wang, J. Wang, and X. Chen, "Secure searchable encryption: A survey," *J. commun. inf. netw.*, vol. 1, no. 4, pp. 52–65, 2016.
[262] Z. Chen, J. Nie, Z. Li, C. Ge, and W. Susilo, "Geometric searchable encryption without false positive and its applications," *Comput. J.*, 2022.
[263] E. Thomas, M. van Deventer, T. Stockhammer, A. C. Begen, M.-L. Champel, and O. Oyman, "Applications and deployments of server and network assisted DASH (SAND)," *Proc. Int. Broadcast. Conv. Conf. (IBC)*, p. 22, 2016.
[264] E. Ohn-Bar, A. Tawari, S. Martin, and M. M. Trivedi, "On surveillance for safety critical events: In-vehicle video networks for predictive driver assistance systems," *Comput. Vis. Image Underst.*, vol. 134, pp. 130–140, 2015.
[265] J. Bhatia, Y. Modi, S. Tanwar, and M. Bhavsar, "Software defined vehicular networks: A comprehensive review," *Int. J. Commun. Syst.*, vol. 32, no. 12, e4005, 2019.
[266] T. Dierks and E. Rescorla, "The Transport Layer Security (TLS) protocol version 1.2," in *Internet Eng. Task Force*, Fremont, CA, USA, RFC 5246, August, 2008.
[267] E. Rescorla, "The Transport Layer Security (TLS) protocol version 1.3," in *Internet Eng. Task Force*, RFC 8446, August, 2018.
[268] T. Ylönen and C. Lonvick, *The Secure Shell (SSH) protocol architecture*, 2006.
[269] A. Fuchs, D. Kern, C. Krauß, and M. Zhdanova, "Securing electric vehicle charging systems through component binding," in *Proc. Int. Conf. Computer Safety, Reliability, and Security (SAFECOMP)*, Springer, 2020, pp. 387–401.
[270] L. Waked, M. Mannan, and A. Youssef, "To intercept or not to intercept: Analyzing TLS interception in network appliances," in *Proc. 2018 Asia Conf. Comput. Commun. Secur.*, 2018, pp. 399–412.
[271] R. Kostromin, "Survey of software configuration management tools of nodes in heterogeneous distributed computing environment.," in *Proc. 2nd Int. Workshop ICCS-DE*, 2020, pp. 156–165.
[272] W. Cabral, C. Valli, L. Sikos, and S. Wakeling, "Review and analysis of cowrie artefacts and their potential to be used deceptively," in *Proc. 2019 Int. Conf. Comp. Sci. Comp. Intell. (CSCI)*, IEEE, 2019, pp. 166–171.
[273] W. Bai, M. Pearson, P. G. Kelley, and M. L. Mazurek, "Improving non-experts' understanding of end-to-end encryption: An exploratory study," in *Proc. IEEE Eur. Symp. Secur. Priv. Wkshp. (EuroS&PW) 2020*, IEEE, 2020, pp. 210–219.

[274] K. Wouters, K. Simoens, D. Lathouwers, and B. Preneel, "Secure and privacy-friendly logging for egovernment services," in *Proc. 3rd Int. Conf. on Availability, Reliability and Security (ARES 2008)*, IEEE, 2008, pp. 1091–1096.

[275] T. Pulls, R. Peeters, and K. Wouters, "Distributed privacy-preserving transparency logging," in *Proc. 12th ACM Wkshp. Priv. Electr. Soc. (WPES)*, 2013, pp. 83–94.

[276] N. Kaaniche and M. Laurent, "BDUA: Blockchain-based data usage auditing," in *Proc. 11th IEEE Int. Conf. CLOUD*, IEEE, 2018, pp. 630–637.

[277] H. Liu, Y. Zhang, S. Zheng, and Y. Li, "Electric vehicle power trading mechanism based on blockchain and smart contract in V2G network," *IEEE Access*, vol. 7, pp. 160 546–160 558, 2019.

[278] M. Baza, M. Nabil, N. Lasla, K. Fidan, M. Mahmoud, and M. Abdallah, "Blockchain-based firmware update scheme tailored for autonomous vehicles," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC) 2019*, IEEE, 2019, pp. 1–7.

[279] E. M. Radi, N. Lasla, S. Bakiras, and M. Mahmoud, "Privacy-preserving electric vehicle charging for peer-to-peer energy trading ecosystems," in *Proc. 2019 IEEE Int. Conf. Commun*, IEEE, 2019, pp. 1–6.

[280] F. Jamil, O. Cheikhrouhou, H. Jamil, A. Koubaa, A. Derhab, and M. A. Ferrag, "PetroBlock: A blockchain-based payment mechanism for fueling smart vehicles," *Appl. Sci.*, vol. 11, no. 7, p. 3055, 2021.

[281] L. Cui, Y. Qu, M. R. Nosouhi, S. Yu, J.-W. Niu, and G. Xie, "Improving data utility through game theory in personalized differential privacy," *J. Comput. Sci. Technol.*, vol. 34, no. 2, pp. 272–286, 2019.

[282] B. Chaudhary and K. Singh, "A blockchain enabled location-privacy preserving scheme for vehicular ad-hoc networks," *Peer-to-Peer Netw. Appl.*, vol. 14, no. 5, pp. 3198–3212, 2021.

[283] Z. Lu, W. Liu, Q. Wang, G. Qu, and Z. Liu, "A privacy-preserving trust model based on blockchain for VANETs," *IEEE Access*, vol. 6, pp. 45 655–45 664, 2018.

[284] A. Boualouache, H. Sedjelmaci, and T. Engel, "Consortium blockchain for cooperative location privacy preservation in 5G-enabled vehicular fog computing," *IEEE Trans. Veh. Technol.*, vol. 70, no. 7, pp. 7087–7102, 2021.

[285] O. Samuel, N. Javaid, F. Shehzad, *et al.*, "Electric vehicles privacy preserving using blockchain in smart community," in *Proc. Int. Conf. Broadband Wireless Comput. Commun. Appl.*, Springer, 2019, pp. 67–80.

[286] R. Liang, B. Li, and X. Song, "Blockchain-based privacy preserving trust management model in VANET," in *Proc. Int. Conf. Adv. Data Mining Appl. (ADMA)*, Springer, 2020, pp. 465–479.

[287] D. Wang, L. Zhang, C. Huang, and X. Shen, "A privacy-preserving trust management system based on blockchain for vehicular networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC) 2021*, IEEE, 2021, pp. 1–6.

[288] G. Yu, X. Zha, X. Wang, *et al.*, "Enabling attribute revocation for fine-grained access control in blockchain-IoT systems," *IEEE Trans. Eng. Manage.*, vol. 67, no. 4, pp. 1213–1230, 2020.

[289] D. J. Currie, C. Q. Peng, D. M. Lyle, B. A. Jameson, and M. S. Frommer, "Stemming the flow: How much can the australian smartphone app help to control COVID-19," *Public Health Res. Pract.*, vol. 30, no. 2, p. 3 022 009, 2020.

[290] K. Aliyev, "An explorative analysis of Azerbaijan's COVID-19 policy response and public opinion," *Cauc. Surv.*, vol. 9, no. 3, pp. 300–319, 2021.

[291] E. Mbunge, "Integrating emerging technologies into COVID-19 contact tracing: Opportunities, challenges and pitfalls," *Diabetes Metab. Syndr.: Clin. Res. Rev.*, vol. 14, no. 6, pp. 1631–1636, 2020.
[292] M. J. Hossain, "Impact of COVID-19 pandemic among health care providers in Bangladesh: A systematic review," *Bangladesh J. Infect. Dis.*, S8–S15, 2020.
[293] P. M. Garrett, Y.-W. Wang, J. P. White, Y. Kashima, S. Dennis, and C.-T. Yang, "High acceptance of COVID-19 tracing technologies in Taiwan: A nationally representative survey analysis," *Int. J. Environ. Res. Publich Health*, vol. 19, no. 6, p. 3323, 2022.
[294] É. Schultz, R. Touzani, J. Mancini, and J. K. Ward, "From contact tracing to COVID-19 pass holder; the tortured journey of the French TousAntiCovid contact tracing app," *Public Health*, vol. 206, pp. 5–7, 2022.
[295] C. Herendy, "How were apps developed during, and for, COVID-19?: An investigation into user needs assessment and testing," in *Proc. 11th IEEE Int. Conf. Cogn. Infocommun. (CogInfoCom) 2020*, IEEE, 2020, pp. 000 503–000 508.
[296] S. Hsaini, H. Bihri, S. Azzouzi, and M. E. H. Charaf, "Contact-tracing approaches to fight COVID-19 pandemic: Limits and ethical challenges," in *Proc. IEEE 2nd Int. Conf. Electron. Control Optim. Comput. Sci. (ICECOCS) 2020*, IEEE, 2020, pp. 1–5.
[297] M. N. Lintvedt, "COVID-19 tracing apps as a legal problem: An investigation of the Norwegian'Smittestopp'App," *Oslo L. Rev.*, vol. 8, p. 69, 2021.
[298] J. Bay, J. Kek, A. Tan, *et al.*, "BlueTrace: A privacy-preserving protocol for community-driven contact tracing across borders," *Government Technology Agency-Singapore, Tech. Rep*, 2020.
[299] S. L. Zhou, X. Jia, S. P. Skinner, W. Yang, and I. Claude, "Lessons on mobile apps for COVID-19 from China," *J. Saf. Sci. Res.*, vol. 2, no. 2, pp. 40–49, 2021.
[300] Y. Lu and L. Zhang, "Social media Wechat infers the development trend of COVID-19," *J. Infect*, vol. 81, no. 1, e82–e83, 2020.
[301] V. Q. Li, L. Ma, and X. Wu, "COVID-19, policy change, and post-pandemic data governance: A case analysis of contact tracing applications in East Asia," *Policy Soc.*, 2022.
[302] T. Sharma, M. M. Islam, A. Das, S. T. Haque, and S. I. Ahmed, "Privacy during pandemic: A global view of privacy practices around COVID-19 apps," in *Proc. ACM SIGCAS Conf. Comput. Sustainable Soc.*, 2021, pp. 215–229.
[303] A. Skiljic, "'Stop COVID-19': The croatian application for contact tracing-overview and privacy-related uncertainties," *Eur. Data Prot. L. Rev.*, vol. 6, p. 433, 2020.
[304] J. Stehlıéková *et al.*, "The corona crisis, data protection and tracking apps in the EU: The Czech and Austrian COVID-19 mobile phone apps in the battle against the virus," *Mezinárodnıé vztahy*, vol. 56, no. 1, pp. 35–67, 2021.
[305] D. J. Leith and S. Farrell, "Contact tracing app privacy: What data is shared by Europe's GAEN contact tracing apps," in *IEEE INFOCOM 2021 - IEEE Conf. Comput. Commun.*, IEEE, 2021, pp. 1–10.
[306] J. Rannikko, P. Tamminen, R. Hellsten, J. P. Nuorti, and J. Syrjänen, "Effectiveness of COVID-19 digital proximity tracing app in Finland," *Clin Microbiol Infect*, 2022.
[307] J. H. Reelfs, O. Hohlfeld, and I. Poese, "Corona-Warn-App: Tracing the start of the official COVID-19 exposure notification app for Germany," in *Proceedings of the SIGCOMM'20 Poster and Demo Sessions*, 2020, pp. 24–26.
[308] T. Alanzi, "A review of mobile applications available in the app and google play stores used during the COVID-19 outbreak," *J. Multidiscip. Healthc.*, vol. 14, p. 45, 2021.

[309] P. Rodrıéguez, S. Graña, E. E. Alvarez-León, *et al.*, "A population-based controlled experiment assessing the epidemiological impact of digital contact tracing," *Nat. Commun.*, vol. 12, no. 1, pp. 1–6, 2021.
[310] S. Geber and T. N. Friemel, "A typology-based approach to tracing-app adoption during the COVID-19 pandemic: The case of the SwissCovid app," *J. Quant. Descr.*, vol. 1, online, 2021.
[311] H. Xu, L. Zhang, O. Onireti, Y. Fang, W. J. Buchanan, and M. A. Imran, "BeepTrace: Blockchain-enabled privacy-preserving contact tracing for COVID-19 pandemic and beyond," *IEEE Internet Things J.*, vol. 8, no. 5, pp. 3915–3929, 2020.
[312] S. Vaudenay, "Analysis of DP3T-between Scylla and Charybdis," Tech. Rep., 2020.
[313] R. L. Rivest, J. Callas, R. Canetti, *et al.*, "The PACT protocol specification," *Private Automated Contact Tracing Team, MIT, Cambridge, MA, USA, Tech. Rep. 0.1*, 2020.
[314] R. Gupta, M. Bedi, P. Goyal, S. Wadhera, and V. Verma, "Analysis of COVID-19 tracking tool in India: Case study of Aarogya Setu mobile application," *Digital Government: Research and Practice*, vol. 1, no. 4, pp. 1–8, 2020.
[315] G. Jung, H. Lee, A. Kim, and U. Lee, "Too much information: Assessing privacy risks of contact trace data disclosure on people with COVID-19 in South Korea," *Public Health Front.*, vol. 8, p. 305, 2020.
[316] J. K. Liu, M. H. Au, T. H. Yuen, *et al.*, "Privacy-preserving COVID-19 contact tracing app: A zero-knowledge proof approach," *Cryptology ePrint Archive*, 2020.
[317] N. D. Shah, E. W. Steyerberg, and D. M. Kent, "Big data and predictive analytics: Recalibrating expectations," *Jama*, vol. 320, no. 1, pp. 27–28, 2018.
[318] A. M. Elbir, G. Gurbilek, B. Soner, A. K. Papazafeiropoulos, P. Kourtessis, and S. Coleri, "Vehicular networks for combating a worldwide pandemic: Preventing the spread of COVID-19," *Smart Health*, vol. 26, p. 100 353, 2022.
[319] H. Stevens and M. B. Haines, "Tracetogether: Pandemic response, democracy, and technology," *Asian Sci. Technol. Soc.*, vol. 14, no. 3, pp. 523–532, 2020.
[320] T. Sharma and M. Bashir, "Use of apps in the COVID-19 response and the loss of privacy protection," *Nat. Med.*, vol. 26, no. 8, pp. 1165–1167, 2020.
[321] H. Cho, D. Ippolito, and Y. W. Yu, "Contact tracing mobile apps for COVID-19: Privacy considerations and related trade-offs," *arXiv preprint arXiv:2003.11511*, 2020.
[322] C. Shachar, J. Engel, and G. Elwyn, "Implications for telehealth in a postpandemic future: Regulatory and privacy issues," *Jama*, vol. 323, no. 23, pp. 2375–2376, 2020.
[323] X. Wang, Z. Ning, M. Zhou, *et al.*, "Privacy-preserving content dissemination for vehicular social networks: Challenges and solutions," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 2, pp. 1314–1345, 2018.
[324] V. K. Yadav, S. Verma, and S. Venkatesan, "Linkable privacy-preserving scheme for location-based services," *IEEE Trans. Intell. Transp. Syst.*, 2021.
[325] H. Zhao, J. Yan, X. Luo, and X. Gua, "Privacy preserving solution for the asynchronous localization of underwater sensor networks," *IEEE/CAA J. Autom. Sin.*, vol. 7, no. 6, pp. 1511–1527, 2020.
[326] Z. Ding, X. Li, C. Jiang, and M. Zhou, "Objectives and state-of-the-art of location-based social network recommender systems," *Acm Comput. Surv. (Csur)*, vol. 51, no. 1, pp. 1–28, 2018.
[327] W. Lin, X. Zhang, L. Qi, *et al.*, "Location-aware service recommendations with privacy-preservation in the Internet of Things," *IEEE Trans. Comput. Soc. Syst.*, vol. 8, no. 1, pp. 227–235, 2020.

[328] R. Al-ani, B. Zhou, Q. Shi, T. Baker, and M. Abdlhamed, "Adjusted location privacy scheme for VANET safety applications," in *NOMS 2020-2020 IEEE/IFIP Netw. Oper. Manag. Symp.*, IEEE, 2020, pp. 1–4.

[329] Y.-B. Zhang, Q.-Y. Zhang, Z.-Y. Li, Y. Yan, and M.-Y. Zhang, "A k-anonymous location privacy protection method of dummy based on geographical semantics.," *Int. J. Netw. Secur.*, vol. 21, no. 6, pp. 937–946, 2019.

[330] I. Ullah, M. A. Shah, A. Wahid, A. Mehmood, and H. Song, "ESOT: A new privacy model for preserving location privacy in Internet of Things," *Telecommun. Syst.*, vol. 67, no. 4, pp. 553–575, 2018.

[331] C. Gao, C. Huang, Y. Yu, H. Wang, Y. Li, and D. Jin, "Privacy-preserving cross-domain location recommendation," *Proc. of the ACM on Intera. Mobile Wear. Ubi. Technol.*, vol. 3, no. 1, pp. 1–21, 2019.

[332] C. A. Ardagna, M. Cremonini, S. D. C. di Vimercati, and P. Samarati, "An obfuscation-based approach for protecting location privacy," *IEEE Trans. Depen. Secu. Comput.*, vol. 8, no. 1, pp. 13–27, 2009.

[333] K. Wei, J. Li, M. Ding, *et al.*, "Federated learning with differential privacy: Algorithms and performance analysis," *IEEE Trans. Inf. Forens. Secu.*, vol. 15, pp. 3454–3469, 2020.

[334] L. Wang, D. Zhang, D. Yang, B. Y. Lim, X. Han, and X. Ma, "Sparse mobile crowdsensing with differential and distortion location privacy," *IEEE Trans. Inf. Forens. Secur.*, vol. 15, pp. 2735–2749, 2020.

[335] F. Koufogiannis, S. Han, and G. J. Pappas, "Optimality of the laplace mechanism in differential privacy," *arXiv preprint arXiv:1504.00065*, 2015.

[336] R. Mendes, M. Cunha, and J. P. Vilela, "Impact of frequency of location reports on the privacy level of geo-indistinguishability," *Proc. Priv. Enh.*, vol. 2020, no. 2, pp. 379–396, 2020.

[337] J. Xiong, R. Ma, L. Chen, *et al.*, "A personalized privacy protection framework for mobile crowdsensing in iiot," *IEEE Trans. Industr. Inform.*, vol. 16, no. 6, pp. 4231–4241, 2019.

[338] H. Zhong, J. Ni, J. Cui, J. Zhang, and L. Liu, "Personalized location privacy protection based on vehicle movement regularity in vehicular networks," *IEEE Syst. J.*, 2021.

[339] Y. Qu, S. Yu, W. Zhou, and Y. Tian, "GAN-driven personalized spatial-temporal private data sharing in cyber-physical social systems," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 4, pp. 2576–2586, 2020.

[340] Y. He, J. Zhang, L. Shuai, J. Luo, X. Yang, and Q. T. Sun, "A personalized secure publishing mechanism of the sensing location data in crowdsensing location-based services," *IEEE Sensors J.*, 2021.

[341] Y. Zheng, Q. Li, Y. Chen, X. Xie, and W.-Y. Ma, "Understanding mobility based on GPS data," in *Proc. of the 10th Int. Conf. Ubi. comput.*, 2008, pp. 312–321.

[342] J. Yuan, Y. Zheng, X. Xie, and G. Sun, "Driving with knowledge from the physical world," in *Proc. of the 17th ACM SIGKDD Int. Conf. Knowl. Disc. data mining*, 2011, pp. 316–324.

[343] V. K. Yadav, S. Verma, and S. Venkatesan, "Efficient and secure location-based services scheme in VANET," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13 567–13 578, 2020.

[344] I. Ullah, M. A. Shah, A. Khan, and G. Jeon, "Privacy-preserving multilevel obfuscation scheme for vehicular network," *Trans. Emerg. Telecommun. Technol.*, e4204, 2020.

[345] S. Gao, M. Zhou, Y. Wang, J. Cheng, H. Yachi, and J. Wang, "Dendritic neuron model with effective learning algorithms for classification, approximation, and prediction," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 2, pp. 601–614, 2018.

[346] P. Zhang, S. Shu, and M. Zhou, "An online fault detection method based on SVM-grid for cloud computing systems," *IEEE/CAA J. Automat. Sinica*, vol. 5, no. 2, pp. 445–456, 2018.
[347] J.-J. Wang and T. Kumbasar, "Optimal PID control of spatial inverted pendulum with Big Bang–Big Crunch optimization," *IEEE/CAA J. Automat. Sinica*, vol. 7, no. 3, pp. 822–832, 2018.
[348] H. Teng, M. Dong, Y. Liu, W. Tian, and X. Liu, "A low-cost physical location discovery scheme for large-scale Internet of Things in smart city through joint use of vehicles and UAVs," *Future Gener. Comput. Syst.*, vol. 118, pp. 310–326, 2021.
[349] L. Sharma, A. Javali, R. Nyamangoudar, R. Priya, P. Mishra, and S. K. Routray, "An update on location based services: Current state and future prospects," in *2017 International Conference on Computing Methodologies and Communication (ICCMC)*, IEEE, 2017, pp. 220–224.
[350] Y. Wang, Z. Cai, X. Tong, Y. Gao, and G. Yin, "Truthful incentive mechanism with location privacy-preserving for mobile crowdsourcing systems," *Comput. Netw.*, vol. 135, pp. 32–43, 2018.
[351] Z. Liu, L. Wu, J. Ke, W. Qu, W. Wang, and H. Wang, "Accountable outsourcing location-based services with privacy preservation," *IEEE Access*, vol. 7, pp. 117 258–117 273, 2019.
[352] Z. Xiong, Z. Cai, Q. Han, A. Alrawais, and W. Li, "Adgan: Protect your location privacy in camera data of auto-driving vehicles," *IEEE Trans. Ind. Inform.*, vol. 17, no. 9, pp. 6200–6210, 2020.
[353] I. A. T. Hashem, V. Chang, N. B. Anuar, *et al.*, "The role of big data in smart city," *Int. J. Inf. Manage.*, vol. 36, no. 5, pp. 748–758, 2016.
[354] C. Dwork, "Differential privacy," vol. 2006, ICALP, 2006, pp. 1–12.
[355] L. Yu, L. Liu, and C. Pu, "Dynamic differential ;ocation privacy with personalized error bounds.," in *Network and Distributed System Security Symposium (NDSS)*, 2017.
[356] R. Shokri, G. Theodorakopoulos, J.-Y. Le Boudec, and J.-P. Hubaux, "Quantifying location privacy," in *2011 IEEE symposium on security and privacy*, IEEE, 2011, pp. 247–262.
[357] Y. Xiao, L. Xiong, S. Zhang, and Y. Cao, "Loclok: Location cloaking with differential privacy via hidden markov model," *Proc. of the VLDB Endowment*, vol. 10, no. 12, pp. 1901–1904, 2017.
[358] C. Chen, D. Zhang, P. S. Castro, N. Li, L. Sun, and S. Li, "Real-time detection of anomalous taxi trajectories from GPS traces," in *International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services*, Springer, 2011, pp. 63–74.
[359] J. Mao, T. Wang, C. Jin, and A. Zhou, "Feature grouping-based outlier detection upon streaming trajectories," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 12, pp. 2696–2709, 2017.
[360] D. Yao, C. Zhang, Z. Zhu, *et al.*, "Learning deep representation for trajectory clustering," *Expert Syst.*, vol. 35, no. 2, e12252, 2018.
[361] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
[362] N. Nikhil and B. Tran Morris, "Convolutional neural network for trajectory prediction," in *Proc. of the European Conference on Computer Vision (ECCV) Workshops*, 2018.
[363] D. Suo, M. E. Renda, and J. Zhao, "Quantifying the tradeoff between cybersecurity and location privacy," *arXiv preprint arXiv:2105.01262*, 2021.

[364] M. Zurbarán, K. Avila, P. Wightman, and M. Fernandez, "Near-rand: Noise-based location obfuscation based on random neighboring points," *IEEE Latin America Trans.*, vol. 13, no. 11, pp. 3661–3667, 2015.

[365] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 770–778.

[366] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[367] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proc. of the IEEE International Conference on Computer Vision (ICCV)*, Oct. 2017, pp. 618–626.

[368] G. Maier, "Openstreetmap, the wikipedia map," *REGION*, vol. 1, no. 1, R3–R10, Dec. 2014. DOI: 10.18335/region.v1i1.70. [Online]. Available: https://openjournals.wu-wien.ac.at/ojs/index.php/region/article/view/70.

[369] M.-h. Oh and G. Iyengar, "Sequential anomaly detection using inverse reinforcement learning," in *Proc. of the 25th ACM SIGKDD International Conference on Knowledge Discovery & data mining*, 2019, pp. 1480–1490.

[370] K. Gray, D. Smolyak, S. Badirli, and G. Mohler, "Coupled igmm-gans for deep multimodal anomaly detection in human mobility data," *arXiv preprint arXiv:1809.02728*, 2018.

[371] Y. Zhao, B. Ma, Z. Wang, Z. Liu, Y. Zeng, and J. Ma, "Trajectory obfuscation and detection in internet-of-vehicles," in *2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, IEEE, 2022, pp. 769–774.

[372] S. Khan, I. Sharma, M. Aslam, M. Z. Khan, and S. Khan, "Security challenges of location privacy in vanets and state-of-the art solutions: A survey," *Future Internet*, vol. 13, no. 4, p. 96, 2021.

[373] S. Oya, C. Troncoso, and F. Pérez-González, "Rethinking location privacy for unknown mobility behaviors," in *2019 IEEE European Symposium on Security and Privacy (EuroS&P)*, IEEE, 2019, pp. 416–431.

[374] Q. Ye, H. Hu, X. Meng, and H. Zheng, "Privkv: Key-value data collection with local differential privacy," in *2019 IEEE Symposium on Security and Privacy (SP)*, IEEE, 2019, pp. 317–331.

[375] S. Narain, A. Ranganathan, and G. Noubir, "Security of gps/ins based on-road location tracking systems," in *2019 IEEE Symposium on Security and Privacy (SP)*, IEEE, 2019, pp. 587–601.

[376] Y. Zhao, J. Zhao, M. Yang, *et al.*, "Local differential privacy-based federated learning for internet of things," *IEEE Internet Things J.*, vol. 8, no. 11, pp. 8836–8853, 2020.

[377] B. Bostanipour and G. Theodorakopoulos, "Joint obfuscation of location and its semantic information for privacy protection," *Comput. Secur.*, vol. 107, p. 102 310, 2021.

[378] A. Pyrgelis, N. Kourtellis, I. Leontiadis, J. Serrà, and C. Soriente, "There goes wally: Anonymously sharing your location gives you away," in *2018 IEEE International Conference on Big Data (Big Data)*, IEEE, 2018, pp. 1218–1227.

[379] P. Asuquo, H. Cruickshank, J. Morley, *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, 2018.

[380] M. S. Sheikh, J. Liang, and W. Wang, "Security and privacy in vehicular ad hoc network and vehicle cloud computing: A survey," *Wireless Communications and Mobile Computing*, vol. 2020, 2020.

[381] M. Yang, B. Ye, Y. Chen, *et al.*, "A trusted de-swinging k-anonymity scheme for location privacy protection," *Journal of Cloud Computing*, vol. 11, no. 1, pp. 1–15, 2022.

[382] G. Li, Q. Zhang, J. Li, J. Wu, and P. Zhang, "Energy-efficient location privacy preserving in vehicular networks using social intimate fogs," *IEEE Access*, vol. 6, pp. 49 801–49 810, 2018.

[383] Y. Jiang, K. Zhang, Y. Qian, and L. Zhou, "Reinforcement-learning-based query optimization in differentially private iot data publishing," *IEEE Internet Things J.*, vol. 8, no. 14, pp. 11 163–11 176, 2021.

[384] X. Li, H. Zhang, Y. Ren, *et al.*, "Papu: Pseudonym swap with provable unlinkability based on differential privacy in vanets," *IEEE Internet Things J.*, vol. 7, no. 12, pp. 11 789–11 802, 2020.

[385] T. Song, N. Capurso, *et al.*, "Enhancing GPS with lane-level navigation to facilitate highway driving," *IEEE Trans. Vel. Technol.*, vol. 66, no. 6, pp. 4579–4591, 2017.

[386] D. Ye, S. Shen, T. Zhu, B. Liu, and W. Zhou, "One parameter defense—defending against data inference attacks via differential privacy," *IEEE Trans. Inf. Forens. Secur.*, vol. 17, pp. 1466–1480, 2022.

[387] Z. Gao, Y. Huang, L. Zheng, H. Lu, B. Wu, and J. Zhang, "Protecting location privacy of users based on trajectory obfuscation in mobile crowdsensing," *IEEE Trans. Industr. Inform.*, vol. 18, no. 9, pp. 6290–6299, 2022. DOI: 10.1109/TII.2022.3146281.

[388] C. Xu, J. Ren, D. Zhang, and Y. Zhang, "Distilling at the edge: A local differential privacy obfuscation framework for iot data analytics," *IEEE Commun. Mag.*, vol. 56, no. 8, pp. 20–25, 2018.

[389] H. Artail and N. Abbani, "A pseudonym management system to achieve anonymity in vehicular ad hoc networks," *IEEE Trans. Dependable Secure Comput.*, vol. 13, no. 1, pp. 106–119, 2015.

[390] R. Yu, J. Kang, *et al.*, "Mixgroup: Accumulative pseudonym exchanging for location privacy enhancement in vehicular social networks," *IEEE Trans. Dependable Secure Comput.*, vol. 13, no. 1, pp. 93–105, 2015.

[391] Z. Liu, Z. Liu, L. Zhang, and X. Lin, "Marp: A distributed mac layer attack resistant pseudonym scheme for vanet," *IEEE Trans. Dependable Secure Comput.*, vol. 17, no. 4, pp. 869–882, 2018.

[392] D. Eckhoff, C. Sommer, *et al.*, "Strong and affordable location privacy in VANETs: Identity diffusion using time-slots and swapping," in *2010 IEEE Vehicular Networking Conference*, IEEE, 2010, pp. 174–181.

[393] K. Sampigethaya, M. Li, L. Huang, and R. Poovendran, "Amoeba: Robust location privacy scheme for vanet," *IEEE Journal on Selected Areas in communications*, vol. 25, no. 8, pp. 1569–1589, 2007.

[394] L. Benarous, B. Kadri, and S. Boudjit, "Alloyed pseudonym change strategy for location privacy in vanets," in *2020 IEEE 17th Annual Consumer Commun. Net. Conf. (CCNC)*, IEEE, 2020, pp. 1–6.

[395] N. Takbiri, A. Houmansadr, D. L. Goeckel, and H. Pishro-Nik, "Matching anonymized and obfuscated time series to users' profiles," *IEEE Trans. Inf. Theory*, vol. 65, no. 2, pp. 724–741, 2018.

[396] H. Al-Balasmeh, M. Singh, and R. Singh, "Framework of data privacy preservation and location obfuscation in vehicular cloud networks," *Concurrency Computat.: Pract. Exper.*, vol. 34, no. 5, e6682, 2022.
[397] W. L. Croft, J.-R. Sack, and W. Shi, "Obfuscation of images via differential privacy: From facial images to general images," *Peer-to-Peer Netw. Appl.*, vol. 14, pp. 1705–1733, 2021.
[398] S. Gopi, Y. T. Lee, and L. Wutschitz, "Numerical composition of differential privacy," *Advances in Neural Information Processing Systems*, vol. 34, pp. 11 631–11 642, 2021.
[399] T. Bao, L. Xu, L. Zhu, L. Wang, and T. Li, "Successive point-of-interest recommendation with personalized local differential privacy," *IEEE Trans. Vel. Technol.*, vol. 70, no. 10, pp. 10 477–10 488, 2021.
[400] Y. Li, Y. Yin, *et al.*, "A secure dynamic mix zone pseudonym changing scheme based on traffic context prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 9492–9505, 2022. DOI: 10.1109/TITS.2021.3125744.
[401] L. Benarous and B. Kadri, "Obfuscation-based location privacy-preserving scheme in cloud-enabled internet of vehicles," *Peer-to-Peer Net. App.*, vol. 15, no. 1, pp. 461–472, 2022.
[402] N. Ahmed, Z. Deng, *et al.*, "A survey on location privacy attacks and prevention deployed with IoT in vehicular networks," *Wireless Commun. Mobile Comput.*, vol. 2022, 2022.
[403] B. Liu, W. Zhou, T. Zhu, L. Gao, T. H. Luan, and H. Zhou, "Silence is golden: Enhancing privacy of location-based services by content broadcasting and active caching in wireless vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 9942–9953, 2016.
[404] W. Li, H. Guo, M. Nejad, and C.-C. Shen, "Privacy-preserving traffic management: A blockchain and zero-knowledge proof inspired approach," *IEEE access*, vol. 8, pp. 181 733–181 743, 2020.
[405] M. Khodaei and P. Papadimitratos, "Cooperative location privacy in vehicular networks: Why simple mix zones are not enough," *IEEE Internet Things J.*, vol. 8, no. 10, pp. 7985–8004, 2020.
[406] M. Dibaei, X. Zheng, *et al.*, "Attacks and defences on intelligent connected vehicles: A survey," *Digital Communications and Networks*, vol. 6, no. 4, pp. 399–421, 2020.
[407] H. Talat and T. a. Nomani, "A survey on location privacy techniques deployed in vehicular networks," in *2019 16th Int. Bhurban Conf. Applied Sci. Technol. (IBCAST)*, IEEE, 2019, pp. 604–613.
[408] J. Liu, C. Zhang, K. Xue, and Y. Fang, "Privacy preservation in multi-cloud secure data fusion for infectious-disease analysis," *IEEE Trans. Mob. Comput.*, vol. early access, 2022.
[409] D. Huang, S. Misra, *et al.*, "Pacp: An efficient pseudonymous authentication-based conditional privacy protocol for vanets," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 3, pp. 736–746, 2011.
[410] M. Luo and Y. Zhou, "An efficient conditional privacy-preserving authentication protocol based on generalized ring signcryption for vanets," *IEEE Trans. Vel. Technol.*, vol. 71, no. 9, pp. 10 001–10 015, 2022. DOI: 10.1109/TVT.2022.3179371.
[411] O. Kanhere and T. S. Rappaport, "Position location for futuristic cellular communications: 5G and beyond," *IEEE Commun. Mag.*, vol. 59, no. 1, pp. 70–75, 2021.
[412] A. Bourdoux, A. N. Barreto, B. van Liempd, *et al.*, "6G white paper on localization and sensing," *arXiv preprint arXiv:2006.01779*, 2020.

[413] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Netw.*, vol. 34, no. 3, pp. 134–142, 2019.

[414] Z. Qin, J. Wang, and Y. Lu, "Monogrnet: A geometric reasoning network for monocular 3D object localization," in *Proc. AAAI Conf. Artif. Intell*, vol. 33, 2019, pp. 8851–8858.

[415] F. Khelifi, A. Bradai, A. Benslimane, P. Rawat, and M. Atri, "A survey of localization systems in Internet of Things," *Mob. Netw. Appl.*, vol. 24, no. 3, pp. 761–785, 2019.

[416] N. A. Azmi, S. Samsul, Y. Yamada, M. F. M. Yakub, M. I. M. Ismail, and R. A. Dziyauddin, "A survey of localization using RSSI and TODA techniques in wireless sensor network: System architecture," in *Proc. 2nd Int. Conf. Telematics Future Gen. Netw. (TAFGEN) 2018*, IEEE, 2018, pp. 131–136.

[417] Y. Wang and K. Ho, "Unified near-field and far-field localization for AoA and hybrid AoA-TDOA positionings," *IEEE Trans. Wirel. Commun.*, vol. 17, no. 2, pp. 1242–1254, 2017.

[418] A. K. Paul and T. Sato, "Localization in wireless sensor networks: A survey on algorithms, measurement techniques, applications and challenges," *J. Sens. Actuator Netw*, vol. 6, no. 4, p. 24, 2017.

[419] J. R. Martinez-de Dios, A. de San Bernabé-Clemente, A. Torres-González, and A. Ollero, "Measurement integration for localization and tracking," in *Cluster-based Localization and Tracking in Ubiquitous Computing Systems*, Springer, 2017, pp. 17–50.

[420] S. Phoemphon, C. So-In, and N. Leelathakul, "A hybrid localization model using node segmentation and improved particle swarm optimization with obstacle-awareness for wireless sensor networks," *Expert Syst. Appl.*, vol. 143, p. 113 044, 2020.

[421] A. A. Zazali, S. K. Subramaniam, and Z. A. Zukarnain, "Flood Control Distance Vector-Hop (FCDV-Hop) localization in wireless sensor networks," *IEEE Access*, vol. 8, pp. 206 592–206 613, 2020.

[422] P. Wang, C. M. Chen, S. Kumari, M. Shojafar, and Y. Liu, "HDMA: Hybrid D2D Message Authentication scheme for 5G-enabled VANET," *IEEE Trans. Intell. Transp. Syst.*, pp. 1–10, 2020.

[423] M. Noura and R. Nordin, "A survey on interference management for Device-to-Device (D2D) communication and its challenges in 5G networks," *Netw. Comput. Appl.*, pp. 130–150, 2016.

[424] W. K. Lai, Y. C. Wang, H. C. Lin, and J. W. Li, "Efficient resource allocation and power control for LTE-A D2D communication with pure D2D model," *IEEE Trans. Veh. Technol.*, vol. PP, no. 99, pp. 1–1, 2020.

[425] W. Y. Al-Rashdan and A. Tahat, "A comparative performance evaluation of machine learning algorithms for fingerprinting based localization in DM-MIMO wireless systems relying on big data techniques," *IEEE Access*, vol. 8, pp. 109 522–109 534, 2020.

[426] D. Shi, X. Zhang, L. Shi, *et al.*, "On improving 5G Internet of Radio Light Security based on LED fingerprint identification method," *Sensors*, vol. 21, no. 4, p. 1515, 2021.

[427] S. Tomic, M. Beko, R. Dinis, M. Tuba, and N. Bacanin, *RSS-AoA-Based Target Localization and Tracking in Wireless Sensor Networks*. River Publishers, 2017.

[428] Y.-Y. Li, G.-Q. Qi, and A.-D. Sheng, "Performance metric on the best achievable accuracy for hybrid ToA/AoA target localization," *IEEE Commun. Letter*, vol. 22, no. 7, pp. 1474–1477, 2018.

[429] B. XU, S. LI, and H. ZHANG, "Low-cost multi-AUV cooperative localization method based on dual-model," *J. Ship Res.*, vol. 15, no. 3, 2020.

[430] J. Lu, X. Chen, M. Luo, and Y. Zhou, "Cooperative localization for multiple AUVs based on the rough estimation of the measurements," *Appl. Soft Comput.*, vol. 91, p. 106 197, 2020.
[431] M. A. Ouameur, M. Caza-Szoka, and D. Massicotte, "Machine learning enabled tools and methods for indoor localization using low power wireless network," *IEEE Internet Things J.*, p. 100 300, 2020.
[432] C. Laoudias, A. Moreira, S. Kim, S. Lee, L. Wirola, and C. Fischione, "A survey of enabling technologies for network localization, tracking, and navigation," *IEEE Commun. Surv. Tutor.*, vol. 20, no. 4, pp. 3607–3644, 2018.
[433] S. Wielandt and L. D. Strycker, "Indoor multipath assisted angle of arrival localization," *Sensors*, vol. 17, no. 11, p. 2522, 2017.
[434] J. Kulmer, "High-accuracy positioning exploiting multipath for reducing the infrastructure," *Ph. D. dissertation, Graz University of Technology*, vol. 3, 2019.
[435] T. Qin, T. Chen, Y. Chen, and Q. Su, "Avp-slam: Semantic visual mapping and localization for autonomous vehicles in the parking lot," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst. IROS 2020*, IEEE, 2020, pp. 5939–5945.
[436] M. J. Piran and D. Y. Suh, "Learning-driven wireless communications, towards 6G," in *Proc. Int. Conf. Comput. Electron. Commun. Eng. (ICCECE) 2019*, IEEE, 2019, pp. 219–224.
[437] A. Chorti, A. N. Barreto, S. Kopsell, *et al.*, "Context-aware security for 6G wireless the role of physical layer security," *arXiv preprint arXiv:2101.01536*, 2021.
[438] H. Viswanathan and P. E. Mogensen, "Communications in the 6G era," *IEEE Access*, vol. 8, pp. 57 063–57 074, 2020.
[439] R. Xie, K. Luo, and T. Jiang, "Joint coverage and localization driven receiver placement in distributed passive radar," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1094–1105, 2020.
[440] C. L. Bamy, F. M. Mbango, D. B. O. Konditi, and P. M. Mpele, "A compact dual-band dolly-shaped antenna with parasitic elements for automotive radar and 5G applications," *Heliyon*, vol. 7, no. 4, e06793, 2021.
[441] F. Wen and H. Wymeersch, "5G synchronization, positioning, and mapping from diffuse multipath," *IEEE Wirel. Commun.*, vol. 10, no. 1, pp. 43–47, 2020.
[442] B. Hu, Z. Shi, and Y. Wang, "Single-sensor based indoor localisation by exploiting multipath propagation," *Electron. Lett.*, vol. 54, no. 3, pp. 179–181, 2018.
[443] L. B. Fertig, M. J. Baden, J. C. Kerce, and D. Sobota, "Localization and tracking with multipath exploitation radar," in *2012 IEEE Radar Conf.*, IEEE, 2012, pp. 1014–1018.
[444] X. Chu, Z. Lu, D. Gesbert, L. Wang, and X. Wen, "Vehicle localization via cooperative channel mapping," *IEEE Trans. Veh. Technol.*, 2021.
[445] G. Schouten, W. Jansen, and J. Steckel, "Simulation of pulse-echo radar for vehicle control and SLAM," *Sensors*, vol. 21, no. 2, p. 523, 2021.
[446] X. Zha, W. Ni, X. Wang, *et al.*, "The impact of link duration on the integrity of distributed mobile networks," *IEEE Trans. Inf. Forensics Secur.*, vol. 13, no. 9, pp. 2240–2255, 2018.
[447] L. Miao, S.-F. Chen, Y.-L. Hsu, and K.-L. Hua, "How does C-V2X help autonomous driving to avoid accidents?" *Sensors*, vol. 22, no. 2, p. 686, 2022.
[448] X. Wang, X. Zha, W. Ni, *et al.*, "Survey on blockchain for Internet of Things," *Comput. Commun.*, vol. 136, pp. 10–29, 2019.
[449] P. Yu, W. Ni, G. Yu, H. Zhang, R. P. Liu, and Q. Wen, "Efficient anonymous data authentication for vehicular ad hoc networks," *Secur. Commun. Netw.*, vol. 2021, 2021.

[450] S. Roth, S. Tomasin, M. Maso, and A. Sezgin, "Localization attack by precoder feedback overhearing in 5G networks and countermeasures," *IEEE Trans. Wirel. Commun.*, vol. 20, no. 7, pp. 4100–4112, 2021. DOI: 10.1109/TWC.2021.3055851.

[451] B. Han, W. Jiang, M. A. Habibi, and H. D. Schotten, "An abstracted survey on 6G: Drivers, requirements, efforts, and enablers," *arXiv preprint arXiv:2101.01062*, 2021.

[452] N. M. Elfatih, M. K. Hasan, Z. Kamal, *et al.*, "Internet of Vehicle's resource management in 5G networks using AI technologies: Current status and trends," *IET Commun.*, vol. 16, no. 5, pp. 400–420, 2022.

[453] X. You, C.-X. Wang, J. Huang, *et al.*, "Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts," *Sci. China Inf. Sci.*, vol. 64, no. 1, pp. 1–74, 2021.

[454] S. S. Al-Bawri, M. S. Islam, H. Y. Wong, L. Lee, and M. T. Islam, "Sub-6 GHz 5G multilayer base station antenna for outdoor localization technique," in *Proc. IEEE Conf. Sustain. Util. Dev. Eng. Technol. CSUDET 2019*, IEEE, 2019, pp. 257–260.

[455] A. S. Abdalla, K. Powell, V. Marojevic, and G. Geraci, "UAV-assisted attack prevention, detection, and recovery of 5G networks," *IEEE Wirel. Commun.*, vol. 27, no. 4, pp. 40–47, 2020.

[456] C. Han, Y. Wang, Y. Li, *et al.*, "Terahertz wireless channels: A holistic survey on measurement, modeling, and analysis," *IEEE Commun. Surv.*, vol. 24, no. 3, pp. 1670–1707, 2022.

[457] K.-T. Feng, L.-H. Shen, C.-Y. Li, *et al.*, "3D on-demand flying mobile communication for millimeter-Wave heterogeneous networks," *IEEE Netw.*, vol. 34, no. 5, pp. 198–204, 2020.

[458] P. Hu, Y. Ma, P. S. Santhalingam, P. H. Pathak, and X. Cheng, "Milliear: Millimeter-Wave acoustic eavesdropping with unconstrained vocabulary," in *IEEE INFOCOM 2022 - IEEE Conf. Comput. Commun.*, IEEE, 2022, pp. 11–20.

[459] C. Wang, F. Lin, T. Liu, *et al.*, "MmEve: Eavesdropping on smartphone's earpiece via COTS mmWave device," in *Proc. 28th Annu. Int. Conf. Mobile Comput. Netw. (MobiCom)*, 2022, pp. 338–351.

[460] K. Ahuja, Y. Jiang, M. Goel, and C. Harrison, "Vid2Doppler: Synthesizing doppler radar data from videos for training privacy-preserving activity recognition," in *Proc. CHI Conf. Hum. Factors Comput. Syst. 2021*, 2021, pp. 1–10.

[461] H. Sarieddeen, N. Saeed, T. Y. Al-Naffouri, and M.-S. Alouini, "Next generation terahertz communications: A rendezvous of sensing, imaging, and localization," *IEEE Commun. Mag.*, vol. 58, no. 5, pp. 69–75, 2020.

[462] Y. Zhu, B. Mao, Y. Kawamoto, and N. Kato, "Intelligent reflecting surface-aided vehicular networks toward 6G: Vision, proposal, and future directions," *IEEE Veh. Technol. Mag.*, vol. 16, no. 4, pp. 48–56, 2021.

[463] I. Rasheed and F. Hu, "Intelligent super-fast vehicle-to-everything 5G communications with predictive switching between mmWave and THz links," *Veh. Commun.*, vol. 27, p. 100 303, 2021.

[464] J. M. Eckhardt, V. Petrov, D. Moltchanov, Y. Koucheryavy, and T. Kürner, "Channel measurements and modeling for low-terahertz band vehicular communications," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1590–1603, 2021.

[465] G. Ananthi and S. Sridevi, "Stacking dilated convolutional autoencoder beamforming for THz wave vehicular ad-hoc networks," *Wirel. Pers. Commun.*, vol. 126, no. 4, pp. 2985–3000, 2022.

[466] C. Han, Y. Wu, Z. Chen, Y. Chen, and G. Wang, "THz ISAC: A physical-layer perspective of terahertz integrated sensing and communication," *arXiv preprint arXiv:2209.03145*, 2022.

[467] B. Ma, X. Lin, X. Wang, *et al.*, "New cloaking region obfuscation for road network-indistinguishability and location privacy," in *Proc. 25th Int. Symp. on Res. in Attacks, Intrusions and Defenses (RAID)*, 2022, pp. 160–170.

[468] G. Singh, A. Srivastava, V. A. Bohara, *et al.*, "Towards 6G-V2X: Hybrid RF-VLC for vehicular networks," *arXiv preprint arXiv:2208.06287*, 2022.

[469] N. Chi, Y. Zhou, Y. Wei, and F. Hu, "Visible light communication in 6G: Advances, challenges, and prospects," *IEEE Veh. Technol. Mag.*, vol. 15, no. 4, pp. 93–102, 2020.

[470] N. H. ElShabasy and S. S. Soliman, "Analysis of multi-hop relaying in VLC-based vehicular networks," in *Proceedings of the 2022 Int. Telecommun. Conf. (ITC-Egypt)*, IEEE, 2022, pp. 1–6.

[471] R. Shaaban, "Visible light communication cyber security vulnerabilities for indoor and outdoor vehicle-to-vehicle communication," Ph.D. dissertation, The University of North Dakota, 2021.

[472] A. Kumar, S. Al-Kuwari, D. N. K. Jayakody, and R. Alkanhel, "Security performance analysis of a NOMA-assisted underwater VLC system under imprecise channel estimations," *IEEE Access*, 2022.

[473] M. Sarfraz, M. F. Sohail, S. Alam, *et al.*, "Capacity optimization of next-generation UAV communication involving non-orthogonal multiple access," *Drones*, vol. 6, no. 9, p. 234, 2022.

[474] R. Cai, Y. Jin, T. Rabczuk, X. Zhuang, and B. Djafari-Rouhani, "Propagation and attenuation of rayleigh and pseudo surface waves in viscoelastic metamaterials," *J. Appl. Phys.*, vol. 129, no. 12, p. 124 903, 2021.

[475] S. Cho, G. Chen, and J. P. Coon, "Cooperative beamforming and jamming for secure VLC system in the presence of active and passive eavesdroppers," *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 4, pp. 1988–1998, 2021.

[476] B. Ma, Z. Liu, Y. Zeng, and J. Ma, "Cooperative jamming for secrecy of wireless communications," in *Proc. Int. Conf. Netw. Netw. Appl. (NaNA) 2018*, IEEE, 2018, pp. 14–21.

[477] S. Zhang, Y. Chang, L. Yan, *et al.*, "Quantum communication networks and trust management: A survey," *Comput. Mater. Continua*, vol. 61, no. 3, pp. 1145–1174, 2019.

[478] P. Sharma, A. Agrawal, V. Bhatia, S. Prakash, and A. K. Mishra, "Quantum key distribution secured optical networks: A survey," *IEEE open j. Commun. Soc.*, vol. 2, pp. 2049–2083, 2021.

[479] A. Singh, K. Dev, H. Siljak, H. D. Joshi, and M. Magarini, "Quantum internet—applications, functionalities, enabling technologies, challenges, and research directions," *IEEE Commun. Surv. Tutor.*, vol. 23, no. 4, pp. 2218–2247, 2021.

[480] M. S. Sharbaf, "Quantum cryptography: A new generation of information technology security system," in *Proc. 6th Int. Conf. Inf. Technol.: New Generations 2009*, IEEE, 2009, pp. 1644–1648.

[481] M. Yarter, G. Uehara, and A. Spanias, "Implementation and analysis of quantum homomorphic encryption," in *Proc. 13th Int. Conf. Int. Conf. Inf. Intell. Syst. Appl.*, IEEE, 2022, pp. 1–5.

[482] S. H. Islam, N. Mishra, S. Biswas, B. Keswani, and S. Zeadally, "An efficient and forward-secure lattice-based searchable encryption scheme for the big-data era," *Comput. Electr. Eng.*, vol. 96, p. 107 533, 2021.

[483] M. Allaix, S. Song, L. Holzbaur, T. Pllaha, M. Hayashi, and C. Hollanti, "On the capacity of quantum private information retrieval from MDS-coded and colluding servers," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 3, pp. 885–898, 2022.

[484] S. Pirandola, J. Eisert, C. Weedbrook, A. Furusawa, and S. L. Braunstein, "Advances in quantum teleportation," *Nat. Photon.*, vol. 9, no. 10, pp. 641–652, 2015.

[485] S. Pirandola and S. Mancini, "Quantum teleportation with continuous variables: A survey," *Laser Phys.*, vol. 16, pp. 1418–1438, 2006.

[486] Z.-Q. He and X. Yuan, "Cascaded channel estimation for large intelligent metasurface assisted massive mimo," *IEEE Wirel. Commun.*, vol. 9, no. 2, pp. 210–214, 2019.

[487] K. Alaaudeen, T. Aruna, and G. Ananthi, "Low-cost, broadband electrically small microstrip antenna using i-shaped metamaterial structure for WLAN/WiMAX and 5G vehicular applications," *Physica B Condens.*, vol. 643, p. 414 144, 2022.

[488] P. Ghate and J. Bredow, "Quasi-optical beamforming approach using vertically oriented dielectric wedges," *Prog. Electromagn. Res. M*, vol. 105, pp. 67–78, 2021.

[489] X. Yuan, Y.-J. A. Zhang, Y. Shi, W. Yan, and H. Liu, "Reconfigurable-intelligent-surface empowered wireless communications: Challenges and opportunities," *IEEE Wirel. Commun.*, vol. 28, no. 2, pp. 136–143, 2021.

[490] E. Basar, "Reconfigurable intelligent surface-based index modulation: A new beyond mimo paradigm for 6G," *IEEE Trans Commun.*, vol. 68, no. 5, pp. 3187–3196, 2020.

[491] Y. Liu, X. Liu, X. Mu, *et al.*, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE Commun. Surv. Tutor.*, vol. 23, no. 3, pp. 1546–1577, 2021.

[492] E. Farsimadan, F. Palmieri, L. Moradi, D. Conte, and B. Paternoster, "Vehicle-to-Everything (V2X) communication scenarios for Vehicular Ad-hoc Networking (VANET): An overview," in *Int. Conf. Comput. Sci. its Appl.*, Springer, 2021, pp. 15–30.

[493] J. Carmona, C. Guindel, F. Garcia, and A. de la Escalera, "eHMI: Review and guidelines for deployment on autonomous vehicles," *Sensors*, vol. 21, no. 9, p. 2912, 2021.

[494] W. Qi, Q. Li, Q. Song, L. Guo, and A. Jamalipour, "Extensive edge intelligence for future vehicular networks in 6G," *IEEE Wirel. Commun.*, vol. 28, no. 4, pp. 128–135, 2021.

[495] J. E. Domeyer, J. D. Lee, H. Toyoda, B. Mehler, and B. Reimer, "Driver-pedestrian perceptual models demonstrate coupling: Implications for vehicle automation," *IEEE Trans. Hum. Mach. Syst.*, 2022.

[496] A. Boukerche, B. Kantarci, and C. Kaptan, "Towards ensuring the reliability and dependability of vehicular crowd-sensing data in GPS-less location tracking," *Pervasive Mob. Comput.*, p. 101 248, 2020.