# Achieving Differentially Private Location Privacy in Edge-assistant Connected Vehicles

Lu Zhou, Le Yu, Suguo Du, Haojin Zhu, *Senior Member, IEEE,* Cailian Chen, *Member, IEEE*

***Abstract*—Connected vehicles can provide safer and more satisfying services for drivers by using information sensing and sharing. However, current network architecture cannot support massive and real-time data transmissions due to the poor-quality wireless links. To provide real-time data processing and improve drivers' security, edge computing is regarded as a promising method to offer more efficient services by placing computing and storage resources at the network edge. In this article, we will introduce the concept of edge-assistant connected vehicles and propose some promising applications to reduce the network traffic and provide real-time services with the help of massive edge nodes. Unlike in the traditional cloud-based connected vehicles, edge nodes are introduced to enable vehicles to obtain real-time and distributed processing services. Furthermore, considering the location privacy issue in the new architecture, we propose a novel differentially privacy-preserving location-based service usage framework deployed on the edge node, designed to provide an adjustable privacy protection solution to balance the utility and privacy. Finally, we conduct extensive experiments to verify the proposed framework.**

***Index Terms*—Connected Vehicles, Edge Computing, Location Privacy, Differential Privacy.**

## I. Introduction

Connected vehicles are emerging as a promising way to improve driving safety and provide innovative services through sharing information about the surrounding environment with other vehicles and infrastructures. According to the expectation of BI Intelligence, 381 million connected vehicles will be on the road by 2020, an increase from 36 million in 2015 [1]. Connected vehicles have attracted the widespread attention of governments and companies including traditional car manufacturers, Internet service providers, and new technology companies. Examples of such programs include the US DOT's Safety Pilot and the EU's eCoMove program.

By equipping vehicles with Internet access and sensing technology, the emerging vehicular networks allow vehicles to share information and driving parameters with other devices, which enables the drivers to be better informed, more coordinated and make the 'smarter' use of vehicular networks. Based on the specific benefits of the connected vehicles, many new applications have been developed, which can be divided into two types. The first type involves the information sharing between vehicles, such as the vehicle platoon, which has great potential to maximize highway throughput [2]. Another one involves the information sharing between a vehicle and remote servers, which can provide more innovative services including navigation, infotainment, and parking [3].

However, the above vehicular applications depend on massive sharing of information in a short period of time. For example, a vehicle needs to request/send considerable data from/to the remote server for decision making, entertainment and other safety-related services in real time in the current vehicular network, which may consume substantial precious wireless bandwidth resources [4], [5] and cannot be achieved by the current vehicular networking due to the poor-quality wireless links. Fortunately, the concept of edge computing in connected vehicles is proposed to address the above challenges.

Edge computing in connected vehicles is regarded as a promising architecture to offer more efficient services by moving some tasks to the edge of the network (e.g., road side infrastructures and base stations) where computing and storage capabilities are available, rather than relying completely on the remote server [6], which can save substantial bandwidth resources. With the help of the edge node, large amounts of data do not need to be transmitted over the poor-quality wireless links. Taking into account the high-speed mobility of vehicles, in [7], the authors proposed a method based on virtual machine migration and transmission power control for minimizing service delays, which can provide efficient transmission and computing services. Considering that most of the vehicular information (e.g., road conditions, parking lot information, and road collision information) is computation-intensive and/or time-sensitive, edge computing is expected to be a promising technology to improve the Quality of Service (QoS) [8] and Quality of the Experience (QoE) of vehicular applications [9]. Therefore, edge computing based data processing fits well with the vehicular networking by offloading the computing-intensive tasks to the edge devices.

In the article, we introduce the Edge-assistant Vehicular Networking architecture (or EVN in short). In this emerging architecture, the cloud layer and vehicular edge nodes form a new network to meet the challenging demands of vehicular applications, which requires the processing of huge amounts of data in real time under the conditions of poor and intermittent connectivity. In the new architecture, massive data will be preprocessed at the edge devices before uploading to the remote server to reduce the network traffic and provide time-sensitive services. In our setting, edge nodes can be deployed within or near the vehicles to provide real-time and distributed processing resources for information aggregation, data forwarding, and authenticity verification. Furthermore, we present

Lu Zhou, Le Yu, and Haojin Zhu are with Department of Computer Science and Engineering, Shanghai Jiao Tong University, China.

Suguo Du is with Department in Antai College of Economics & Management, Shanghai Jiao Tong University, China.

Cailian Chen is with Department of Automation, Shanghai Jiao Tong University, China.

Corresponding author: Haojin Zhu (email: zhu-hj@sjtu.edu.cn).

some potential applications for the EVN, including information confusion applications and location-based subscription applications. Compared with the cloud-based vehicular applications, the proposed applications can provide computation-intensive and/or time-sensitive services.

In addition to the above benefits, the edge-assistant architecture can also provide a more practical location privacy-preserving scheme compared with the traditional architecture. Under the traditional architecture, differential privacy (DP) is a widely used method for providing a strong privacy guarantee. Previous works have proposed many mechanisms [10], [11], including the advanced approaches to refine the query results from the noisy location by extending the query radius and filtering the retrieved results. However, this method will incur more overhead since the client needs to filter many useless results. If we deploy the privacy-preserving mechanism on the edge node and offload the task to the edge node, The efficiency of privacy-preserving mechanism will be significantly improved.

Furthermore, current methods [10], [11] use the strong privacy notion $\epsilon$-differential privacy to ensure location privacy. As the coverage of an edge node is small, strong privacy guarantee will cause the noisy location to be farther away from the coverage, which has a great impact on the service quality. Therefore, we use an alternative notion of privacy, $(\epsilon, \delta)$-differential privacy [12] to guarantee location privacy, which means the privacy loss is bound by $\epsilon$ with a probability of at least $1 - \delta$. In the new notion, service quality can be guaranteed by loosing the probability $\delta$.

Considering the above issues, we design a novel differentially location privacy-preserving framework that can ensure location privacy in the coverage of an edge node with $(\epsilon, \delta)$-differential privacy and deploy this new mechanism on the edge node. When a vehicle uses the location-based service, it first submits its location to the edge node, and then the edge node executes the privacy-preserving mechanism for querying the POI and filtering the useless results. The proposed framework is mainly composed of two parts: Differentially Privacy-preserving Location-based Service Usage (Pri-LBS) mechanism and Privacy Level Adjustment (PLA) module, which are designed to enable vehicles to request useful information based on the submitted location without revealing their location privacy and to identify a tradeoff between privacy and service quality by providing an adjustable privacy level.

The contributions of this paper are summarized as follows:

- We introduce a novel Edge-assistant Vehicular Networking architecture (or EVN in short) and present some promising applications that use this architecture.
- We analyze the novel location privacy issue in this new architecture and define the location privacy model for this location issue.
- We propose a novel differentially privacy-preserving location-based service usage framework, which can protect individual location privacy while ensuring that service providers provide normal services.
- We give the detailed experiments to validate the proposed scheme. Results have demonstrated the privacy-preserving framework is robust against honest-but-curious attackers.

The rest of this paper is organized as follows. Firstly, we introduce the related work in Section II. We then briefly introduce the novel Edge-assistant Vehicular Networking architecture and some promising applications in Section III and Section IV, respectively. Section V presents the unique location security and privacy challenges under this emerging architecture. Section VI gives the personalized differential privacy-preserving location-based service usage framework. In Section VII, we conduct the detailed experiment. Finally, we conclude the paper in Section VIII.

## II. Related Work

In this section, we will present some existing works about edge computing and its role in promoting the development of connected vehicles. Then, we will introduce the definition of differential privacy and its mechanisms in privacy protection.

### A. Edge Computing in Connected Vehicles

Currently, edge computing has caused wide attention from both academia and industry. By allowing the computations to be executed near the data sources, edge computing can address concerns that cannot be satisfied by the cloud computing, including latency, bandwidth costs, security and privacy. In general, an edge node is any computing or networking resource (e.g., smartphone, roadside unit) between the cloud and the end devices (e.g., vehicles) [6], [13], [14]. In [15], the authors gave a detailed introduction from the perspective of the architecture, promising applications and the comparison between edge computing and cloud computing. This work also pointed out its advantages in reducing latency, bandwidth, and increasing security. Based on the basic architecture, many researchers are investigating the use of edge computing in various scenarios, including Google's SDN-based Internet peering edge routing infrastructure [16], online job dispatching and scheduling [17], and Edge Fabric [18].

Among the massive research areas, edge computing in connected vehicles is regarded as a promising field for deployment. Some existing works [19], [20], [21] discussed the possibility of introducing edge computing into connected vehicles and designed efficient task scheduling and control algorithms. The authors pointed out that edge computing can provide real-time data processing for vehicles, which cannot be achieved by the current cloud-based model. In [22], the authors proposed an information-centric mobile edge computing model for connected vehicles. However, these works only discussed the possibility of introducing edge computing in connected vehicles. Different from the previous works, we propose two important and practical applications for this architecture from the perspective of information gathering and distribution. More importantly, we present the novel location privacy-preserving service usage framework for edge-assistant connected vehicles, which allows vehicles to use normal services without compromising their location privacy.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2018.2876419, IEEE Internet of Things Journal



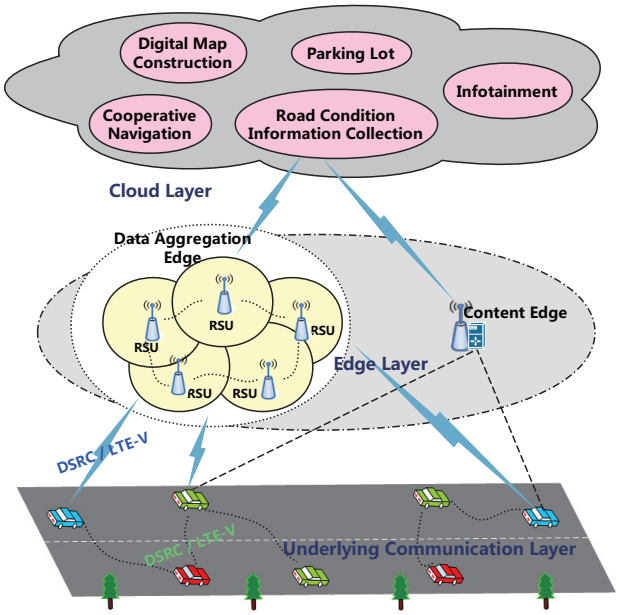



Fig. 1: The architecture of the Edge-assistant Vehicular Networking

### B. Differential Privacy

Differential privacy is proposed by Cynthia Dwork [23]; it is a classic privacy protection notion in the field of statistical databases. In the special scenario, attackers can execute two neighboring queries with a difference of one record. Using the neighboring queries, individual information can be leaked, which is coined as the differential attack. Generally, differential privacy can provide the privacy protection between neighboring databases that differ only in one row. Differential privacy mechanism can ensure that adding/deleting a record has a negligible effect on the query result. The typical approach to realize this privacy guarantee is to add random noise fitting a special distribution to the result, such as Laplace distribution. Based on the basic definition, differential privacy has been used in many applications, such as privacy-preserving data aggregation [24] and private stream processing [25].

How to protect locations with differential privacy is an important field for the researchers from both of the academia and the industry. In [26], the authors discussed the possibility of designing privacy-preserving location pattern mining in a database and designed a quadtree spatial decomposition technique to fit differential privacy. In [10], the authors firstly proposed the scheme to protect a single location with differential privacy and defined a formal notion of privacy, geo-indistinguishability, for location-based systems. Xiao et al. [11] further discussed this issue and proposed a systematic solution to protect location privacy with differential privacy by considering temporal correlations of a moving user's locations.

## III. Architecture of Edge-assistant Connected Vehicles

In this section, we present a classic architecture of the EVN as shown in Fig. 1. The conventional vehicular ad hoc network (or VANET) allows the information transmission in a vehicle-to-vehicle (or V2V), or vehicle-to-roadside-unit (or V2R) fashion. In the emerging EVN, the network architecture is comprised of the following three layers:

- ***Cloud Layer:*** In this architecture, the cloud has the same functions as the traditional architecture. The cloud is expected to support a wide range of infotainment and road-safety related applications, such as digital map construction, parking lots, and so on.
- ***Edge Layer:*** In the edge layer, considering the large amount of the data shared between vehicles and servers, it is desirable to deploy some edge nodes that are in the proximity of the vehicles to accelerate the data processing. The RSUs (Roadside Units) or base stations are good candidates for "edge nodes", which can be deployed along the road-side to serve as the intermediate node between the cloud and vehicles to accelerate the data aggregation and distribution processes. In addition of the basic data storage and forwarding functions, edge nodes can be used to accelerate the road condition information collection and distribution processes between cloud servers and vehicles by data preprocessing, which can provide the acceleration function for supporting real-time computational and storage requirements.
- ***Underlying Communication Layer:*** In the client layer, vehicles are equipped with the On-Board-Unit (OBU) to communicate with other devices. Dedicated Short Range Communication (or DSRC) and LTE-V are two typical technologies that enable the future V2X communications. To support V2X communications, IEEE 802.11p (commonly called "DSRC") is proposed for vehicle-to-vehicle and vehicle-to-infrastructure communications [27]. In China, long-term evolution-V (LTE-V) has been proposed as a systematic and integrated V2X solution based on time-division LTE (TD-LTE) [28]. Different from DSRC, LTE-V is expected to leverage the existing LTE infrastructure and the spectrum to improve the communication distance among vehicles for providing the longer response time in case of emergency.

In summary, by leveraging the benefits of edge computing, EVN is expected to play an important role in the future vehicular network by providing a wide range of new applications, which are illustrated as follows.

## IV. The Promising Applications in EVN

In this section, we present some potential applications for EVN, including the information confusion applications and location-based subscription applications.

### A. Crowdsourcing-based Information Confusion Applications

Crowdsourcing is an important application of vehicular networking. By being equipped with the processing devices, communication devices, and a series of sensing devices (e.g., chemical detectors, still/video cameras, vibration and acoustic sensors), vehicles can be used to collect the important auxiliary information such as digital map updating, point of interest (POI), road traffic information, and traffic collision information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2018.2876419, IEEE Internet of Things Journal



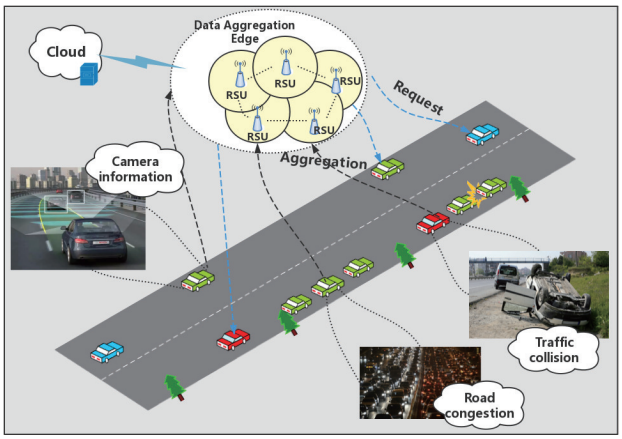



Fig. 2: Real-Time Data Aggregation

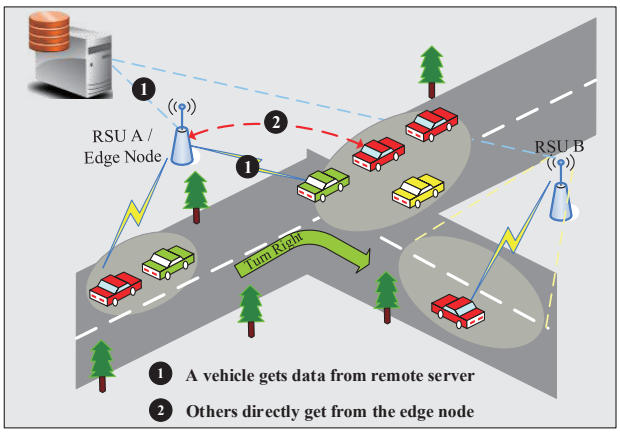



Fig. 3: Location-based Information Request

Traditionally, they are collected by massive staff and require the expensive sensing tools (e.g., specialized cars, sensing coil, and road side cameras), which may consume a lot of manpower and resources. However, due to the high updating frequency of the ground information (e.g., roads, POI, and traffic lights) and resource limitations, the traditional collection method cannot meet the high demand of real-time environment sensing. Therefore, crowdsourcing is regarded as a promising method for data aggregation [29]. As shown in Fig.2, the sensors on vehicles can provide real-time and massive sensing data to the service providers using crowdsourcing, which can update the sensing information in a real-time manner and thus significantly enhance the sensing accuracy at a reduced cost. Note that the collected sensing data can be aggregated and processed at the edge nodes and then propagated to the nearby vehicles, which can significantly reduce the transmission overhead and propagation delay.

**Featured Application: Crowdsourced Digital Map Construction** Digital maps are typically generated via satellite images and/or manual entry by trained personnel, which cannot timely cope with the dynamic changes of the physical world and thus fail to keep the maps up-to-date. Driving recorders equipped with high-precision cameras are the common devices in today's vehicles and will generate massive video streams about the roads, which can be used for map construction. The latest artificial techniques can be adopted to recognize the images in the video and learn the road signs, which can significantly speed up the process of map construction. The video streams from different vehicles can be collected and locally aggregated by the edge nodes, which then forward the aggregated results (including road signs, road shape, and POI) to the cloud. Crowdsensing-based models can meet the market's demand for digital maps because of their considerable benefits with respect to cost and efficiency.

### *B. Location-based Subscription Applications*

From the perspective of massive vehicular applications, location is an important parameter when drivers subscribe to the desired data, including POI, road conditions, and traffic collision information. Therefore, location-based services are important in current vehicular networking. In addition, vehicles within a certain proximity may query the same information related to this location in a very short time interval. However, under the current vehicular networking scheme, vehicles in the proximity should request multiple copies of the data from remote servers, which occupies the considerable wireless bandwidth resources and cannot achieve the requirement of real-time processing in the emergency situations.

As shown in Fig.3, in the edge-enabled architecture, a vehicle can subscribe to the desired data from the remote server by setting its location as the interest for data searching, which can be stored in the edge node in a short time. After that, other vehicles can retrieve the desired data from the surrounding edges that have a copy of the data, not just from the original producer of the data. With the deployment of the edge-assistant networking, there is great potential to reduce data latency and network traffic for providing efficient and real-time services to vehicles. Therefore, edge-enabled architecture matches the vehicular network better.

**Featured Application: Edge-enabled Augmented Reality (AR) Vehicle Navigation** AR vehicle navigation is a promising application in the future of vehicular networking that can provide more accurate and human-like services than current navigation systems by clearly pointing out where the driver will go, and highlighting traffic signs and POI based on what the driver can actually see. AR scene construction needs massive data in real time, which cannot be supported by the current vehicular network. Near a vehicle, other vehicles that are using the AR vehicle navigation will request similar data. In the edge-assistant AR vehicle navigation, these vehicles can cooperate and obtain data from the edge node, which can support timely AR vehicle navigation.

## V. Edge Computing in Providing Location Privacy Protection

In this section, we focus on the location privacy issue of the edge-assistant vehicular network. Similar to the location-based services in the traditional architecture, the location privacy issue also exists in the new architecture. Differential privacy is the most used method to provide robust location privacy

guarantees that are not affected by any auxiliary information available to the attacker. Current methods [10], [11] proposed to ensure the location privacy use the notion of differential privacy with Laplacian noise, which are deployed in the client. As pointed out in [10], the noisy location generated by the privacy-preserving mechanism will affect the accuracy of query results. Therefore, the authors improved the mechanism by extending the query results and filtering the retrieved results at the client. However, filtering many useless results will incur massive overhead, which will increase the load of the client.

Therefore, it is reasonable to offload this task to the edge node which is assumed to be trusted to reduce the cost. As the edge node is close to the user side, it can provide a more secure environment than the cloud server. Even if the edge server is untrusted, we can create an independent environment that can only be controlled by the user and deploy the hardware protection technologies (such as ARM TrustZone, Intel SGX [30]) to provide a secure environment for users.

In this section, we discuss how to design a location privacy-preserving mechanism that can ensure the differential privacy in the coverage of the edge node. First, we give the new location privacy model: We assume that the attacker is an honest-but-curious service provider, which means it is honest in executing the protocol to provide the location-based services, but curious in inferring users' private information based on the collected data. We assume that $r$ is the radius of the coverage (circle $S$) of the edge node, and as a prior knowledge, the attacker learns the vehicle is within the coverage (circle $S$) of an edge node. We deploy the the privacy protection mechanism $\mathcal{M}$ to protect users' privacy by generating the noisy location. Our goal is to guarantee that the attacker can only learns that the vehicle is within the coverage (circle $S$) and cannot know the real location.

Current methods [10], [11] use the strong privacy notion, $\epsilon$-differential privacy, to ensure the location privacy. However, this strong privacy guarantee will have a great impact on the service quality, which is not suitable in the edge-assistant architecture. As the coverage of an edge node is small (such as 5G base station, in which the radius is generally approximately 0.2km [31]), some location-sensitive applications (e.g., find the nearest edge node for executing computing tasks) will have a great deviation. To solve this problem, in this paper, we use a relaxed notion of privacy, $(\epsilon, \delta)$-differential privacy [12], which means the privacy loss is bound by $\epsilon$ with a probability of at least $1 - \delta$.

Therefore, to ensure the above guarantee, we design a novel differentially privacy-preserving framework at the edge node by introducing the notion of $(\epsilon, \delta)$-differential privacy, coined as Pri-EVN, which can protect individual location privacy while ensuring that service providers provide normal services. At the same time, the proposed location privacy mechanism can minimize the impact on the service quality as much as possible. When a vehicle uses the location-based service, it first submit its location to the edge node for launching the proposed location privacy mechanism, and then the edge node executes the differentially privacy-preserving mechanism for querying the desired results (such as POI) from the service providers and filtering the useless results according to the real location. In the next section, we will introduce the novel differentially privacy-preserving mechanism which fits the edge-assistant architecture.

## VI. Differentially Privacy-preserving Location-based Service Usage Framework

To thwart the location privacy threats mentioned above and provide a location privacy guarantee for vehicles, we propose a differentially privacy-preserving location-based service usage framework, which is coined as Pri-EVN. As shown in Fig. 4, the proposed framework is mainly composed of two parts: Differentially Privacy-preserving Location-based Service Usage (Pri-LBS) mechanism and Privacy Level Adjustment (PLA) module, which are designed to enable vehicles to request useful information based on the submitted location without revealing their privacy and identify the tradeoff between the privacy and the service quality by providing the adjustable privacy level.

### A. The Proposed Pri-LBS Mechanism to Thwart the Location Leakage Attack

In this section, we propose a novel differentially privacy-preserving location-based service usage (Pri-LBS) mechanism to protect vehicles' location privacy with a strict privacy bound. The basic idea of Pri-LBS is to send the actual location to the edge node to generate the noisy location and then send the noisy location to the service provider for querying and at the same time ensure that the attacker cannot infer vehicles' real location based on the submitted location. We use a new notion of privacy, $(\epsilon, \delta)$-differential privacy [12] to protect the location privacy. Then, we discuss how to add a new distribution, Gaussian distribution, into the differentially privacy-preserving service usage mechanism.

*1) Differential Privacy Definition for Location Privacy Under the Edge-assistant Architecture:* As shown in section V, the attacker is assumed to be honest-but-curious and is interested in collecting vehicles' location privacy. As a prior knowledge, the attacker learns the vehicle is within the coverage (circle $S$) of an edge node. Our goal is to guarantee that the attacker can only learns that the vehicle is within the coverage (circle $S$) and cannot know the real location. Based on the concept of differential privacy [10], [32], we give the formal definition of location privacy, $(r_1, \epsilon, \delta)$-Geo-indistinguishability, which can ensure the privacy loss bounded by $\epsilon$ with a probability of at least $1 - \delta$.

**Definition 1.** *($(r_1, \epsilon, \delta)$-Geo-indistinguishability) Assuming $l_1$ is the real location of the vehicle and $l_2$ is any location in the circle $S_1$ (the center is $l_1$) with a radius $r_1$, a mechanism $\mathcal{M}$ satisfies $(r_1, \epsilon, \delta)$-geo-indistinguishability iff for any submitted location $l_0$:*

$$P(l_0|l_1) \leq e^{\epsilon} P(l_0|l_2) + \delta \qquad \forall l_1, l_2 : d(l_1, l_2) \leq r_1,\ r_1 = 2r \tag{1}$$

which means two location $l_1$ and $l_2$ produce a submitted location with similar probabilities. Based on the submitted location $l_0$, the attacker does not have a non-negligible probability of differentiating which one is the real location. This

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2018.2876419, IEEE Internet of Things Journal



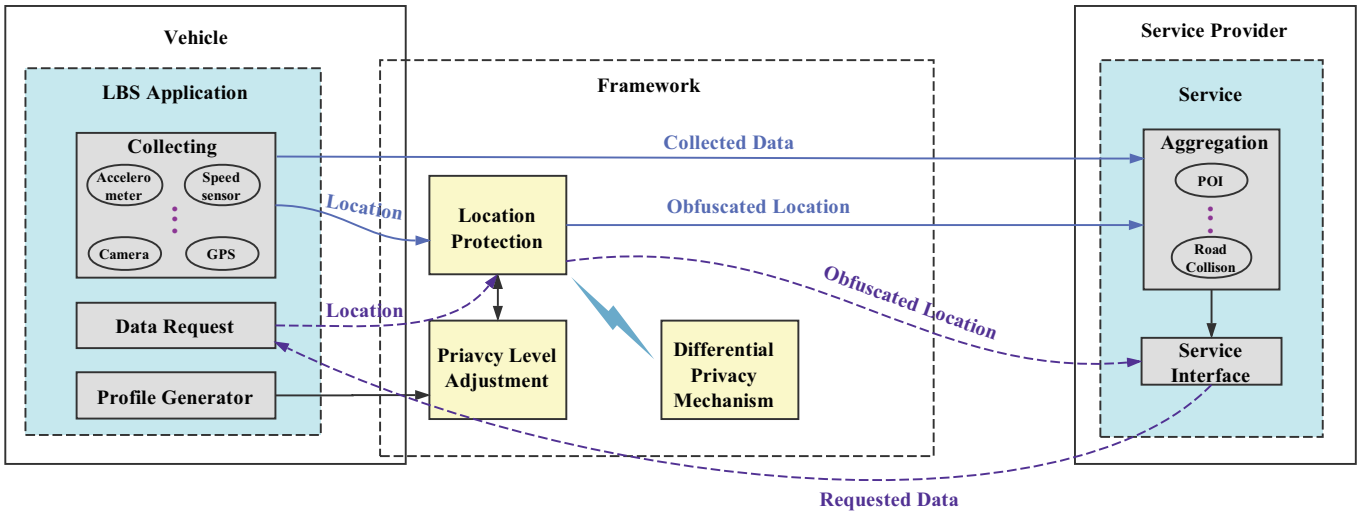



Fig. 4: Privacy-preserving Location Usage Framework

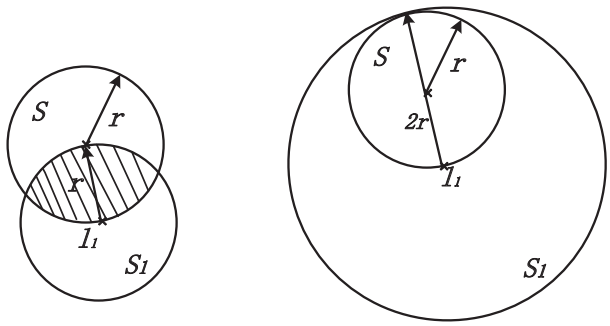



Fig. 5: Ensuring the differential privacy in different ranges

definition restricts the advantage that an adversary can get from observing the obfuscated location $l_0$.

Note that we should ensure the differential privacy in a circle with a radius of $r_1 = 2r$, rather than $r$. We prove its reasonableness by the Fig. 5. We consider the extreme circumstance that the real location $l_1$ is located in the circle, as shown in Fig. 5. If we ensure the $(r, \epsilon, \delta)$-Geo-indistinguishability (Fig. 5 (left)), only part of the circle $S$ can meet the privacy guarantee (the shadowed part). Therefore, it is reasonable to achieve $(r_1, \epsilon, \delta)$-Geo-indistinguishability.

*2) A Mechanism for $(r_1, \epsilon, \delta)$-Geo-indistinguishability:* In this section, we will propose the concrete method to add the noise which can fit the $(r_1, \epsilon, \delta)$-Geo-indistinguishability. We propose a scheme to generate the noisy location which follows two dimensional Gaussian Distribution.

The probability density function of two dimensional Gaussian distribution is given by

$$f(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{2}$$

where $\sigma$ is standard deviation. Then we give the the privacy loss which is bounded by $\epsilon$ to obtain $\sigma$ which fits the $(r_1, \epsilon, \delta)$-Geo-indistinguishability. The privacy loss $Pri_{loss}$ is defined as follows:

$$\begin{aligned} Pri_{loss} &= \left| \ln \frac{e^{-\frac{x^2+y^2}{2\sigma^2}}}{e^{-\frac{(x-\Delta x)^2+(y-\Delta y)^2}{2\sigma^2}}} \right| \\ &= \left| -\frac{x^2+y^2}{2\sigma^2} + \frac{(x-\Delta x)^2+(y-\Delta y)^2}{2\sigma^2} \right| \\ &= \left| \frac{-(x*\Delta x + y*\Delta y) + \frac{(\Delta x)^2+(\Delta y)^2}{2}}{\sigma^2} \right| \\ &\leq \epsilon \end{aligned} \tag{3}$$

Which can be converted into:

$$\begin{aligned} & \left| \frac{-(x*\Delta x + y*\Delta y) + \frac{(\Delta x)^2+(\Delta y)^2}{2}}{\sigma^2} \right| \\ & \leq |x*\Delta x + y*\Delta y| + \frac{(\Delta x)^2+(\Delta y)^2}{2} \\ & \leq \epsilon\sigma^2 \end{aligned} \tag{4}$$

According to the Cauchy-Buniakowsky-Schwarz and $\sqrt{(\Delta x)^2+(\Delta y)^2} \leq r_1$, we can obtain:

$$\sqrt{x^2+y^2} \leq \frac{\epsilon\sigma^2}{r_1} - \frac{r_1}{2} \tag{5}$$

As defined by the $(r_1, \epsilon, \delta)$-Geo-indistinguishability, we ensure the privacy loss bounded by $\epsilon$ with probability at least $1-\delta$, which means:

$$P[\sqrt{x^2+y^2} \geq \frac{\epsilon\sigma^2}{r_1} - \frac{r_1}{2}] < \delta \tag{6}$$

$$\sigma \geq \frac{r_1}{\epsilon}\sqrt{\ln\frac{1}{\delta^2} + \epsilon} \tag{7}$$

The Guassian nosie with parameter $\sigma \geq \frac{r_1}{\epsilon}\sqrt{\ln\frac{1}{\delta^2} + \epsilon}$ is $(r_1, \epsilon, \delta)$-Geo-indistinguishability.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2018.2876419, IEEE Internet of Things Journal



*3) Obfuscation with Gaussian Noise Fitting $(r_1, \epsilon, \delta)$-Geo-indistinguishability:* In this section, we explore how to add the Gaussian noise fitting $(r_1, \epsilon, \delta)$-Geo-indistinguishability to an actual location. The basic idea is that we shift the location by adding the noise generated randomly from the following two dimensional Gaussian distribution:

$$f(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \qquad \sigma \geq \frac{r_1}{\epsilon}\sqrt{\ln\frac{1}{\delta^2}+\epsilon} \quad (8)$$

which can be transformed into the following form under the polar coordinates:

$$f(r',\theta) = \frac{r'}{2\pi\sigma^2} e^{-\frac{r'^2}{2\sigma^2}} \quad (9)$$

where $r'$ is the distance between the actual location $l_1$ and the noisy location $l_0$, and $\theta$ is the angle that indicates the direction of the line $l_1 l_0$. The joint probability density of $r'$ and $\theta$ are described as follows:

$$f(r') = \int_0^{2\pi} \frac{r'}{2\pi\sigma^2} e^{-\frac{r'^2}{2\sigma^2}} \, d\theta = \frac{r'}{\sigma^2} e^{-\frac{r'^2}{2\sigma^2}} \quad (10)$$

$$f(\theta) = \int_0^{\infty} \frac{r'}{2\pi\sigma^2} e^{-\frac{r'^2}{2\sigma^2}} \, dr' = \frac{1}{2\pi} \quad (11)$$

From the above equations we can see that these two variables are independent, which means we can draw them separately.

Assuming the actual location is $l_1 = (s,t)$, the noisy location that follows two dimensional Gaussian distribution can be generated as

$$l_0 = (s + r'\cos\theta, t + r'\sin\theta) \quad (12)$$

*4) Optimal Privacy Mechanism Under the Tolerable Service Quality Loss:* By deploying the differential privacy protection mechanism $\mathcal{M}$, we can ensure the attacker cannot infer the actual location even if the attacker has the prior knowledge that the vehicles' location must be within the coverage of the edge node. However, the noisy location must have the impact on the service quality. It is reasonable to obtain the maximum privacy level under the tolerable service quality loss. In this section, we formalize the problem as an optimization problem between the privacy and the service quality.

First, we define the service quality based on the utility notion which is introduced in [10]. The service quality can be represented by the tolerable maximum offset distance $D$ and the probability $\gamma$:

**Definition 2.** *($(D,\gamma)$-Service Quality) Let d(,) denote the Euclidean metric and $l_1$ be the actual location of the vehicle, the differential privacy protection mechanism $\mathcal{M}$ satisfies ($(D,\gamma)$-Service Quality) if:*

$$P[d(l_1, \mathcal{M}(l_1)) \geq D] \leq \gamma \quad (13)$$

Note that $d(l_1, \mathcal{M}(l_1)) = r'$. Therefore, the probability can be computed by the cumulative distribution function (cdf) $F(r')$:

$$F(r') = 1 - e^{-\frac{r'^2}{2\sigma^2}} \quad (14)$$

$$P[d(l_1, \mathcal{M}(l_1)) \geq D] = 1 - F(D) = e^{-\frac{D^2}{2\sigma^2}} \leq \gamma \quad (15)$$

We assume the tolerable service quality loss is $(D', \gamma')$, which means the minimum service quality that the vehicle can tolerate. This constraints the privacy level, which can be used as a threshold for the mechanism $\mathcal{M}$.

Therefore, the optimization problem (find the optimal privacy level) under the given service quality could be described as follows:

$$\textbf{Minimize} \quad Pri_{loss} \quad (16)$$

$$\textbf{subject to}$$

$$P[d(l_1, \mathcal{M}(l_1)) \geq D'] = e^{-\frac{D'^2}{2\sigma^2}} \leq \gamma' \quad (17)$$

$$\sigma \geq \frac{r_1}{\epsilon}\sqrt{\ln\frac{1}{\delta^2}+\epsilon} \quad (18)$$

$$Pri_{loss} \leq \epsilon \quad (19)$$

From the optimization problem, we can get the optimal two tuples $(\epsilon^*, \delta^*)$ (ensuring $P(pri_{loss} \leq \epsilon^*) \geq 1 - \delta^*$) under the specified $(D', \gamma')$, which has the following relationship:

$$\epsilon^* = \sqrt{\frac{2\ln\gamma'({r_1}^2 \ln\frac{1}{{\delta^*}^2} - 2{r_1}^4 \ln\gamma')}{-{D'}^2}} - \frac{{r_1}^2 \ln\gamma'}{{D'}^2} \quad (20)$$

Similarly, we can set the optimization problem for finding the optimal service quality $(D^*, \gamma^*)$ (ensuring $P[d(l_1, \mathcal{M}(l_1)) \geq D^*] \leq \gamma^*$) under the given privacy loss bound $(\epsilon', \delta')$ with the same constraint conditions:

$$D^* = \frac{r_1}{\epsilon'}\sqrt{-2\ln\gamma^*(\ln\frac{1}{{\delta'}^2} + \epsilon')} \quad (21)$$

### B. Privacy Level Adjustment (PLA)

Note that, under different situations users have different privacy requirements. For example, when people stays at their top locations (e.g., home and office) where they always go, they want to protect these sensitive locations with a higher level of privacy at the expense of the service quality. However, when they are in public regions, it is not reasonable to sacrifice so many service quality to protect their location privacy. Therefore, it is reasonable to adjust the privacy level $\frac{1}{\epsilon}$ which is bounded by the $(\epsilon, \delta)$ to obtain the better service quality. This requirement can be achieved by the context-based privacy protection scheme proposed in [33], which can automatically learn users' privacy preferences based on the users' habits by using a decision tree model.

## VII. Evaluation

In this section, we demonstrate the effectiveness and efficiency of the proposed differentially privacy-preserving location-based service usage mechanism by evaluating its accuracy on different applications. Then, we explain its efficiency under different parameters with respect to privacy level and service quality.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2018.2876419, IEEE Internet of Things Journal



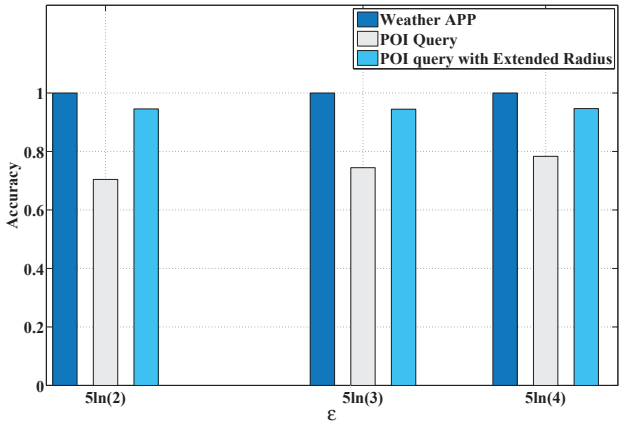



Fig. 6: The impact of the deployment of the framework on different applications

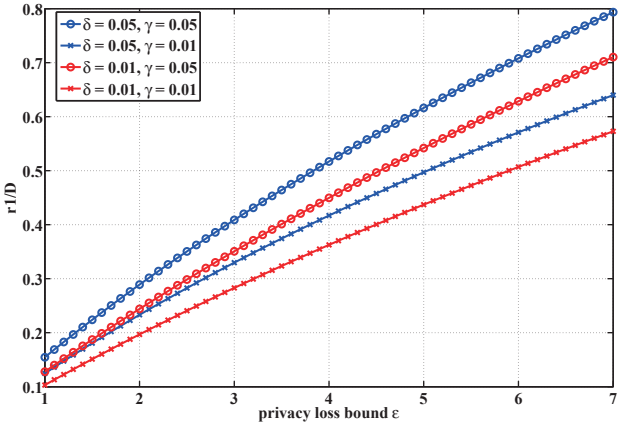



Fig. 7: The relationship between $\epsilon$ and $\frac{r1}{D}$

### A. The Effectiveness of Pri-EVN

We consider the proposed location-based subscription applications in the edge-assistant vehicular networking scenario. In our setting, drivers communicate with the honest-but-curious service provider via an edge node. In this paper, we consider two types of applications which have the different requirements for location accuracy: low-location-sensitive applications and high-location-sensitive applications. We implement the proposed privacy-preserving mechanism using Java, including random parameters generation, coordinate transformation, and the result analysis.

First, we evaluate the accuracy of the proposed framework when using the different types of applications. To assess the impact of our framework on different scenarios, we deploy the differential privacy-preserving location-based service usage mechanism on two typical vehicular applications: Weather APP and POI query. Note that, we use the API provided by the Baidu Map for POI query. We evaluate how much accuracy will be reduced when deploying our mechanism with different privacy loss with bound $\epsilon$. In particular, the accuracy refers to the difference between the query results by these two locations (actual location and the noisy location) and we use multiple (e.g., 10) query results as the average accuracy to eliminate the error. We choose the different $\epsilon$ as used in [10], and the result between the accuracy and $\epsilon$ under different applications is shown in Fig. 6. In this figure, we can see that our mechanism has no effect on low-location-sensitive applications. For the high-location-sensitive applications, the accuracy decreases by approximately $25\%$ when deploying our differential privacy-preserving location-based service usage mechanism. Therefore, it is better to use the alternative method: increasing the query radius to ensure the accuracy. We know the maximum offset ratio under the specific $\epsilon$, therefore we can extend the query radius by the maximum offset distance $D$. Note that, the offset can be eliminated by filtering the results. The accuracy under this query radius is shown by the third bar (POI query with extended radius).

Then, we evaluate the relationship between the privacy loss bound $\epsilon$ (where $\epsilon$ is the privacy loss bound) and service quality $\frac{r_1}{D}$, where $r_1$ is the radius and $D$ is the maximum offset distance with the privacy protection mechanism $\mathcal{M}$ ($D$ is bounded by $P[d(l_1, \mathcal{M}(l_1)) \geq D] \leq \gamma$). We evaluate the relationship with very small parameters $\delta$ and $\gamma$, which means a very strong bound in the range of $D$ and $\epsilon$. The result under different parameters ($\delta = 0.01, 0.05$ and $\gamma = 0.01, 0.05$) is shown in Fig. 7. As shown in this figure, the privacy loss bound $\epsilon$ and the service quality $\frac{r_1}{D}$ have the linear relationship, which means with the increase of privacy protection level, the service quality will not be drop too fast. Therefore, under different scenarios as described in the previous section, drivers could adjust the privacy level to find the tradeoff with the service quality.

### B. The Efficiency Under Different Parameters

As described in the previous subsection, different applications have different requirements with respect to the privacy level and the service quality. However, because the differential privacy-preserving mechanism has the probabilistic nature, it is impossible to guarantee the privacy level and the service quality with $100\%$ probability. In this section, we further evaluate the impact of parameters $(\delta, \gamma)$ on privacy level and the service quality under the optimal situation, and discuss how to adjust them to fit different scenarios.

First, we illustrate the relationship of $(\epsilon, \delta)$ under the given service quality, which is shown in Fig. 8. From this figure we have the following insight: with the decrease of $1 - \delta$ (the confidence of the privacy loss bound), the privacy loss bound $\epsilon$ will decrease under the given service quality. Therefore, we can slightly increase $\delta$ to get the higher privacy level $\frac{1}{\epsilon}$ when deploying our mechanism on applications which do not need the strong confidence. For example, when deploying the mechanism on the users' traces, attackers cannot recover the trace even if a small portion of the locations has the risk of leakage.

Fig. 9 shows the relationship of $D$ and $\gamma$ under the specified privacy level and the fixed $r_1 = 0.2$ (the same parameter as used in [10]). When $\epsilon = 5\ln(2), \delta = 0.01$, the offset distance can be controlled within $500m$ with a probability of 0.95, within $440m$ with a probability of 0.9. Similar to the privacy level, we can also adjust $\gamma$ to obtain the higher service quality.

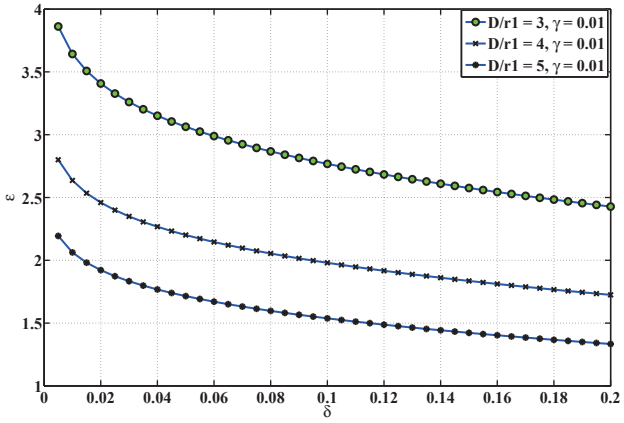



Fig. 8: The relationship between $\epsilon$ and $\delta$ under the given service quality

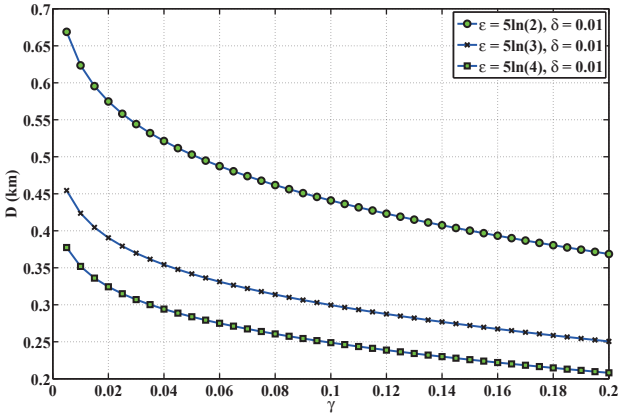



Fig. 9: The relationship between $D$ and $\gamma$ under the given privacy level

For example, in some location-based news push applications, users do not need the strict service quality guarantee, so we can choose a relatively large value.

## VIII. Conclusion

In this article, we propose an edge-assistant networking in connected vehicles including its architecture, advantages, and some promising applications. Then, we discuss the location security and privacy requirements for the new paradigm and some possible solutions. In addition, we propose a novel differentially privacy-preserving service usage framework to achieve these security and privacy requirements. Taking into account the limited sizes of the caching in vehicles, our future work will consider the problem of how to address the location- and time-dependent cached data in a fast-moving environment.

## Acknowledgment

This work was supported in part by National Science Foundation of China under Grant U1405251, 61672350, 61622307, U1401253, in part by Shanghai Science and Technology Committee, China, under Grant 18511111502, and in part by the China Scholarship Council (201806230109).

## References

[1] (2016) Automotive industry trends: Iot connected smart cars vehicles. [Online]. Available: http://www.businessinsider.com/internet-of-things-connected-smart-cars-2016-10

[2] D. Jia, K. Lu, J. Wang, X. Zhang, and X. Shen, "A survey on platoon-based vehicular cyber-physical systems," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 263–284, 2017.

[3] Z. Ning, F. Xia, N. Ullah, X. Kong, and X. Hu, "Vehicular social networks: Enabling smart mobility," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 16–55, 2017.

[4] D. Zhang, Z. Chen, M. K. Awad, N. Zhang, H. Zhou, and X. S. Shen, "Utility-optimal resource management and allocation algorithm for energy harvesting cognitive radio sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 12, pp. 3552–3565, 2016.

[5] D. Zhang, Y. Qiao, L. She, R. Shen, J. Ren, and Y. Zhang, "Two time-scale resource management for green internet of things networks," *IEEE Internet of Things Journal*, 2018.

[6] W. Shi and S. Dustdar, "The promise of edge computing," *Computer*, vol. 49, no. 5, pp. 78–81, 2016.

[7] T. G. Rodrigues, K. Suto, H. Nishiyama, and N. Kato, "Hybrid method for minimizing service delay in edge cloud computing through vm migration and transmission power control," *IEEE Transactions on Computers*, vol. 66, no. 5, pp. 810–819, 2017.

[8] D. Zhang, R. Shen, J. Ren, and Y. Zhang, "Delay-optimal proactive service framework for block-stream as a service," *IEEE Wireless Communications Letters*, 2018.

[9] J. Ren, Y. Guo, D. Zhang, Q. Liu, and Y. Zhang, "Distributed and efficient object detection in edge computing: Challenges and solutions," *IEEE Network*, vol. PP, no. 99, pp. 1–7, 2018.

[10] M. E. Andrés, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, "Geo-indistinguishability: Differential privacy for location-based systems," in *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*. ACM, 2013, pp. 901–914.

[11] Y. Xiao and L. Xiong, "Protecting locations with differential privacy under temporal correlations," in *ACM Sigsac Conference on Computer and Communications Security*, 2015, pp. 1298–1309.

[12] C. Dwork, K. Kenthapadi, F. Mcsherry, I. Mironov, and M. Naor, "Our data, ourselves: Privacy via distributed noise generation," in *Advances in Cryptology - EUROCRYPT 2006, International Conference on the Theory and Applications of Cryptographic Techniques, St. Petersburg, Russia, May 28 - June 1, 2006, Proceedings*, 2006, pp. 486–503.

[13] X. Peng, J. Ren, L. She, D. Zhang, J. Li, and Y. Zhang, "Boat: A block-streaming app execution scheme for lightweight iot devices," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1816–1829, 2018.

[14] J. Ren, H. Guo, C. Xu, and Y. Zhang, "Serving at the edge: A scalable iot architecture based on transparent computing," *IEEE Network*, vol. 31, no. 5, pp. 96–105, 2017.

[15] M. Satyanarayanan, "The emergence of edge computing," *Computer*, vol. 50, no. 1, pp. 30–39, 2017.

[16] K.-K. Yap, M. Motiwala, J. Rahe, S. Padgett, M. Holliman, G. Baldus, M. Hines, T. Kim, A. Narayanan, A. Jain *et al.*, "Taking the edge off with espresso: Scale, reliability and programmability for global internet peering," in *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*. ACM, 2017, pp. 432–445.

[17] H. Tan, Z. Han, X.-Y. Li, and F. C. Lau, "Online job dispatching and scheduling in edge-clouds," in *INFOCOM 2017-IEEE Conference on Computer Communications, IEEE*. IEEE, 2017, pp. 1–9.

[18] B. Schlinker, H. Kim, T. Cui, E. Katz-Bassett, H. V. Madhyastha, I. Cunha, J. Quinn, S. Hasan, P. Lapukhov, and H. Zeng, "Engineering egress with edge fabric: steering oceans of content to the world," in *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*. ACM, 2017, pp. 418–431.

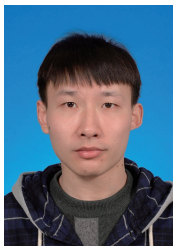
[19] T. S. Darwish and K. A. Bakar, "Fog based intelligent transportation big data analytics in the internet of vehicles environment: Motivations, architecture, challenges, and critical issues," *IEEE Access*, vol. 6, pp. 15 679–15 701, 2018.

[20] Y. Lai, F. Yang, L. Zhang, and Z. Lin, "Distributed public vehicle system based on fog nodes and vehicular sensing," *IEEE Access*, vol. 6, pp. 22 011–22 024, 2018.

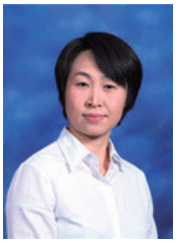
[21] C.-M. Huang, M.-S. Chiang, D.-T. Dao, W.-L. Su, S. Xu, and H. Zhou, "V2v data offloading for cellular network based on the software defined network (sdn) inside mobile edge computing (mec) architecture," *IEEE Access*, vol. 6, pp. 17 741–17 755, 2018.

[22] D. Grewe, M. Wagner, M. Arumaithurai, I. Psaras, and D. Kutscher, "Information-centric mobile edge computing for connected vehicle environments: Challenges and research directions," in *Proceedings of the Workshop on Mobile Edge Communications*. ACM, 2017, pp. 7–12.
[23] C. Dwork, "Differential privacy: A survey of results," in *International Conference on Theory and Applications of MODELS of Computation*, 2008, pp. 1–19.
[24] E. Shi, T. H. H. Chan, and E. Rieffel, "Privacy-preserving aggregation of time-series data," *Annual Network & Distributed System Security Symposium*, 2011.
[25] Y. Chen, A. Machanavajjhala, M. Hay, and G. Miklau, "Pegasus: Data-adaptive differentially private stream processing," in *ACM Sigsac Conference*, 2017, pp. 1375–1388.
[26] S. S. Ho and S. Ruan, "Differential privacy for location pattern mining," in *ACM Sigspatial International Workshop on Security and Privacy in Gis and Lbs*, 2011, pp. 17–24.
[27] Y. L. Morgan, "Notes on dsrc & wave standards suite: Its architecture, design, and characteristics," *IEEE Communications Surveys & Tutorials*, vol. 12, no. 4, pp. 504–518, 2010.
[28] (2015) 3gpp, tr 22.891 v. 2.0.0, feasibility study on new services and markets technology enablers. [Online]. Available: http://www.3gpp.org/DynaReport/22891.htm
[29] Z. He, J. Cao, and X. Liu, "High quality participant recruitment in vehicle-based crowdsourcing using predictable mobility," in *Computer Communications*, 2015, pp. 2542–2550.
[30] R. Pettersen, H. D. Johansen, and D. Johansen, "Secure edge computing with arm trustzone," in *Proc. 2nd Int. Conf,. Internet Things, Big Data Secur.(IoTBDS)*, 2017, pp. 102–109.
[31] A. I. Sulyman, A. T. Nassar, M. K. Samimi, G. R. MacCartney, T. S. Rappaport, and A. Alsanie, "Radio propagation path loss models for 5g cellular networks in the 28 ghz and 38 ghz millimeter-wave bands," *IEEE Communications Magazine*, vol. 52, no. 9, pp. 78–86, 2014.
[32] C. Dwork and A. Roth, *The Algorithmic Foundations of Differential Privacy*. Now Publishers Inc., 2014.
[33] H. Li, H. Zhu, S. Du, X. Liang, and X. Shen, "Privacy leakage of location sharing in mobile social networks: Attacks and defense," *IEEE Transactions on Dependable and Secure Computing*, 2016.

**Lu Zhou** received the B.Eng degree in Computer Science and Technology from Sichuan University, China, in 2015. He is currently pursuing the Ph.D. degree in Computer Science and Technology at Shanghai Jiao Tong University, China. His research interests include security and privacy protection in Connected Vehicles and Cognitive Radio Networks.



**Le Yu** received the B.Eng degree in Computer Science and Technology from Shanghai Jiao Tong University, China, in 2014. He is currently pursuing the Ph.D. degree in Computer Science and Technology at Shanghai Jiao Tong University, China. His research interests include location privacy, differential privacy.



**Suguo Du** is an Associate Professor in Department of Management Science, Shanghai Jiao Tong University, China. She received her PhD degree in School of Mathematical and Information Sciences from Coventry University, U.K., in 2002. Her current research interests include Risk and Reliability Assessment, Vehicular Networks Security and Privacy Protection and Social Networks Security Management. Her research work has been supported by National Science Foundation of China.

**Haojin Zhu** (IEEE M'09-SM'16) received his B.Sc. degree (2002) from Wuhan University (China), his M.Sc. degree (2005) from Shanghai Jiao Tong University (China), both in computer science and the Ph.D. in Electrical and Computer Engineering from the University of Waterloo (Canada), in 2009. Since 2017, he has been a full professor with Computer Science department in Shanghai Jiao Tong University. His current research interests include network security and privacy enhancing technologies. He published more than 40 international journal papers, including JSAC, TDSC, TPDS, TMC, TWC, TVT, and 60 international conference papers, including ACM CCS, ACM MOBICOM, ACM MOBIHOC, IEEE INFOCOM, IEEE ICDCS. He received a number of awards including: IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award (2014), Top 100 Most Cited Chinese Papers Published in International Journals (2014), Supervisor of Shanghai Excellent Master Thesis Award (2014), Distinguished Member of the IEEE INFOCOM Technical Program Committee (2015), Outstanding Youth Post Expert Award for Shanghai Jiao Tong University (2014), SMC Young Research Award of Shanghai Jiao Tong University (2011). He was a co-recipient of best paper awards of IEEE ICC (2007) and Chinacom (2008), IEEE GLOBECOM Best Paper Nomination (2014), WASA Best Paper Runner-up Award (2017). He received Young Scholar Award of Changjiang Scholar Program by Ministry of Education of P.R. China in 2016.

**Cailian Chen** (IEEE S'03-M'06) received the B.Eng. and M.Eng. degrees in Automatic Control from Yanshan University, P. R. China in 2000 and 2002, respectively, and the Ph.D. degree in Control and Systems from City University of Hong Kong, Hong Kong SAR in 2006. She joined Department of Automation, Shanghai Jiao Tong University in 2008 as an Associate Professor. She is now a Full Professor. Before that, she was a senior research associate in City University of Hong Kong (2006) and postdoctoral research associate in University of Manchester, U. K. (2006-2008). She was a Visiting Professor in University of Waterloo, Canada (2013-2014). Prof. Chen has worked actively on various topics such as wireless sensor networks and industrial applications, computational intelligence and distributed situation awareness, cognitive radio networks and system design, Internet of Vehicles and applications in intelligent transportation, and distributed optimization. She has authored and/or coauthored 2 research monographs and over 100 referred international journal and conference papers. She is the inventor of more than 20 patents. Dr. Chen received the prestigious "IEEE Transactions on Fuzzy Systems Outstanding Paper Award" in 2008, and "Best Paper Award of The Ninth Int. Conference on Wireless Communications and Signal Processing" in 2017. She won the First Prize of Natural Science Award twice from The Ministry of Education of China in 2006 and 2016, respectively. She was honored "Changjiang Young Scholar" by Ministry of Education of China in 2015 and "Excellent Young Researcher" by NSF of China in 2016. Prof. Chen has been actively involved in various professional services. She serves as Associate Editor of IEEE TVT, PPNA (Springer), The World Scientific Journal: Computer Science, and ISRN Sensor Networks. She also served as Guest Editor of IEEE TVT, Symposium TPC Co-chair of IEEE Globecom 2016 and VTC2016-fall, Workshop Co-chair of WiOpt'18, and TPC member of many flagship conferences including IEEE Globecom, IEEE ICC, IEEE VTC, ICCVE and IEEE WCCI.