IRDP Ride system: A Privacy Preservation System for Online Ride-hailing

Lei Zhang, Shiyi Lin, Chao Wang, Jing Li, Yi Liu, Yue Sun

***Abstract*—Online ride-hailing and navigation (ORHN) is a significant application in Location-based services (LBS). With this application, users can get convenient services in daily travel. However, during the process of ORHN, the user has to confront the violation of personal privacy (such as the initial point, the destination, and the moving path) from the taxi driver or the ride-hailing sharer. Therefore, in order to cope with the problem of privacy violation, in this paper, an initial, route, and destination preservation ride system (short for IRDP Ride system) is proposed. In the IRDP Ride system, the initial point, the destination, and the route can be preserved during the process of ORHN, no matter whether the adversary (such as the un-trusted driver or sharer) utilizes the analysis method in the same car (such as the range attack or the segment direction attack). In this paper, before elaborating on the IRDP Ride system, two types of methods used for analyzing the initial point and destination were given. Then based on the characteristics of Voronoi and the navigation deviation, three algorithms used in this system for providing privacy preservation service are proposed. In the last part of this paper, the IRDP Ride system's security and stability are discussed from the theoretical analysis. At the same time, this system is also verified with several groups of experiments in the real environment to determine the optimization parameters, and then this system is also compared with other similar systems to demonstrate its superiority.**

***Index Terms*—Privacy preservation, Location-based services, Voronoi, navigation deviation**

I. Introduction

Nowadays, along with the development of position and wireless technology as well as the prosperity of the smartphone, Location-based services (LBS) with the platform of the smartphone has become an important component of people's daily life [1]. In this service, the user just provides its location and some others require information and then get the corresponding service in return (such as navigation, points of interest (PoIs) searching, advertisement pushing as well as online ride-hailing), so it can provide a great convenience for people's daily travel and daily life. However, during the whole process of getting this service, as the user has to provide the precise location, the personal privacy of the user will inevitably confront the issues of privacy violation. In addition, as the privacy violation will also bring about embarrassing situations or some other inconvenience and even the incident of personal security, users will pay more attention to their privacy. As a result, a lot of corresponding algorithms and systems have been proposed to cope with the problem of privacy violation.

Among these algorithms and systems, they can be briefly classified into two categories: privacy preservation for requirement and privacy preservation for service. In the category of privacy preservation for the requirement, the target is to provide preservation service for the user's requirement information. There are a number of algorithms and systems in this category, and they usually assume that the user wants to search for a specified target and provides the precise location and requirement to the LBS provider, and then the user gets results from the LBS provider once or continuously. Under this model, the usually utilized strategy is *k*-anonymity [2], which requires the user or the central server to provide at least *k* similar locations or requirements to the LBS provider to generalize the real requirement. Currently, algorithms and systems are concentrated on the model of differential privacy in long-term observation attacks [3], the query generalization with blockchain models [4], the distributed privacy preservation for mix-context [5] as well as the metric for attack and privacy preservation in query service [6] and so on. As the query information can be adjusted before sending it to the LBS server, these algorithms and systems usually consider the requirement as the main body and just generalize or obfuscate this information to achieve the target of privacy preservation.

Different from the category of privacy preservation for the requirement, in the category of privacy preservation for service, the system provides privacy preservation service for the service (such as a coupon or ride-hailing [7]). In fact, this category is a passive service, as the service provider no longer provides service for the user according to the query, but for the service, so the strategy used in the category of privacy preservation for the requirement is useless for privacy preservation for service. Therefore, the algorithms and systems are concentrated on the privacy-preserving fair meeting location and a group of nearest neighbors searching [8,

This work was supported in part by the Natural Science Foundation of Heilongjiang Province of China under Grant No. LH2021F054, the Excellent Discipline Team Project of Jiamusi University under Grant No.JDXKTD-2019008, the Basic Scientific Research Operating Expenses of Heilongjiang Provincial Universities and Colleges for Excellent Innovation Team under name (The research and development team of privacy theory and technology), the Basic Scientific Research Operating Expenses of Heilongjiang Provincial Universities and Colleges under Grant No.2020-KYYWF-0225.*(Corresponding author: Chao Wang*; *Jing Li.)*

Lei Zhang, Shiyi Lin, Chao Wang, Jing Li, Yi Liu, Yue Sun are with the College of Information science and Electronic Technology, Jiamusi University, Jiamusi, Heilongjiang, 154007, (e-mail: zhanglei@jmsu.edu.cn).

9] as well as on traffic monitoring and smart cities [10, 11], etc. In addition, the strategy for location privacy in camera data of auto-driving and the strategy of privacy-preserving for covid-19 contact tracing[12, 13] is also concerned.

However, these algorithms or systems are mainly focused on privacy preservation in the process of sharing with other users [14] or in the process of tracing [15]. Just a few of them are concerned about privacy in the whole process of ORHN, such as systems for protecting the initial point [16], the moving trajectory [17] as well as the ride car and the sharer [18, 19], etc. In addition, some algorithms or systems mainly focused on utilizing the cryptographic primitives to securely and efficiently estimate the shortest distances between riders and drivers in road networks approximately [20], or utilize the Minhash method to filter out dissimilar routes in advance and reduce computational costs and communication overheads [21] or utilize property-preserving hash with road network embedding to support privacy-preserving ride-matching services [22]. But do not consider the sharer and taxi driver may also be the adversary to jeopardize the user's sensitive information during the whole process of ORHN. Thus, in this paper, we proposed a novel system to provide privacy preservation for users in the ORHN, which we called the initial, route, and destination preservation ride system short for the IRDP Ride system. With this system, the initial point, the destination as well as the route of the user can be preserved. Then the contribution of this paper can be summarized as follows.

1） We proposed two types of methods used for analyzing the initial point and destination and gave the model of them as metrics to verify the privacy violation in the process of ORHN.

2） We proposed the IRDP Ride system and three algorithms used in this system are given to protect the initial point, the destination as well as the route simultaneously. This system is based on the characteristics of Voronoi and navigating deviation.

3） We discussed the security and stability of the IRDP Ride system with theoretical analysis so as to maintain the usability of this system, and then the test results of engineering applications in the real environment were given.

4） We discussed the parameters setting and compared the IRDP Ride system with other similar systems to demonstrate its superiority.

According to the contribution, we organize the remainder of this paper as follows. In Section II, we present the privacy problems, the basic idea, the system architectures, and the definitions. Section III presents the details of the IRDP Ride system, and then we discuss three algorithms used in this system. In Section IV we develop the details to achieve system security and stability, based on the theoretical analysis. In Section IV, we present engineering applications, optimization parameters and discuss the simulation results with other similar systems. Finally, we give more related work similar to the IRDP Ride system in Section V, and we conclude our work as well as discuss further work in Section VI.

II. Definitions and System model

A. Privacy problems

In the process of ORHN, two entities in the same car called the taxi driver and the sharer can be seen as the un-trusted adversary. Both of them will threaten the user's privacy such as the destination and the route, and the taxi driver will also threaten the user's privacy at the initial point. In order to get the above-mentioned information of the user, they can utilize the range attack and segments direction attack to infer the user's privacy.

Definition 1 (Range attack). In general, a user usually chooses the closest place to the initial point or destination as the pick-up or the drop-off point. If the pick-up or the drop-off point can be denoted as l, an adversary can utilize this point as the center and gradually expand the circle to calculate the real point. Suppose that there are k PoIs in a range and denoted as $P=\{p_1, p_2, ..., p_k\}$, if the distance between the location l and a special point p_s in this set satisfies $d(l, p_s) < d(l, p_i)$, where $i \neq s \in k$, then the point p_s will be the initial point or destination.

Definition 2 (Segment direction attack). If the route that the user utilizes in ORHN can be divided into several segments and denoted as $T=\{T_{s1}, T_{s2}, ..., T_{sn}\}$. Then for a special point D_s in the potential set of destinations $D=\{D_1, D_2, ..., D_k\}$, if this point satisfies $\sum_{j=1}^{n}(T_{sj} \to D_s) > \sum_{j=1}^{n}(T_{sj} \to D_i)$, where $i \neq s \in k$, the special point D_s can be seen as the destination.

B. The basic idea and conception

As the adversary can utilize the range attack and segments direction attack to infer the user's privacy, the desired conception to cope with the above two attacks is to break the condition that the adversary can utilize in these attacks. For example, if the adversary cannot get or get wrong $d(l, p_s) < d(l, p_i)$ and $\sum_{j=1}^{n}(T_{sj} \to D_s) > \sum_{j=1}^{n}(T_{sj} \to D_i)$, then conditions for the range attack and segments direction attack are broken.

For the range attack, as the adversary wants to get $d(l, p_s) < d(l, p_i)$ to choose the potential point of initial or destination, the best way is to find a feasible pick-up or a feasible drop-off point that satisfies the following two conditions. The first condition is that the distance $d(\cdot)$ between the real l and p_s is no longer the shortest. The second condition is that once the real l and p_s is the shortest, there will exist several similar points that the adversary difficult to distinguish. Then based on the conception of finding such a location l, the basic idea to realize the above two conditions is to make the PoI in the annulus with an

uncertain radius r and at least k similar points also located in this annulus.

For the segments direction attack, as the adversary calculates $\sum_{j=1}^{n}(T_{sj} \to D_s) > \sum_{j=1}^{n}(T_{sj} \to D_i)$ for the set of destinations, the target of privacy preservation is to shift the direction of each segment to mislead the uniform of segments. Then each segment will have the same probability to correlating various destinations and satisfied $\sum_{j=1}^{n}(T_{sj} \to D_1) = \sum_{j=1}^{n}(T_{sj} \to D_2) = ... = \sum_{j=1}^{n}(T_{sj} \to D_k)$, i.e. each destination in the set has the same probability to be the real destination for the segments direction attack.

C. System architecture

As mentioned in the previous section, based on the basic idea of privacy preservation and the feature of two entities in ORHN, the architecture of the IRDP Ride system can be the centralization architecture, i.e. the IRDP Ride system is deployed in a central server, and provides privacy preservation and navigation service for the user and the taxi driver. The architecture of deployment can be seen in Fig.1.

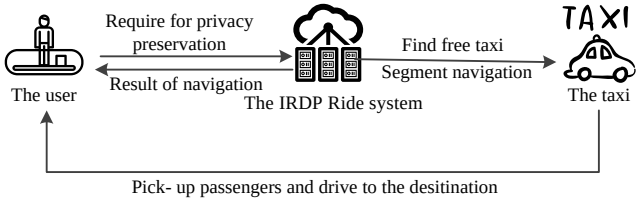


Fig. 1. The architecture of the IRDP Ride system.

In fact, besides the centralization architecture, the IRDP Ride system can also be deployed in hybrid architecture or distributed architecture. In these two architectures, the user conducts the PoI information with the mobile device and sends the navigation information to the taxi driver once the car is in a new Voronoi grid. No matter what types of architecture the user selects, the taxi driver and the car sharer can still be considered the potential adversary and can utilize the range attack and segments direction attack to infer the user's privacy. Therefore, the target of the IRDP Ride system is to ensure that the driver and the sharer are difficult to infer the real information about the user such as the initial point, the destination as well as the routing.

D. Preliminaries

In order to achieve the target of privacy preservation in ORHN, the IRDP Ride system has to find two types of methods to break the conditions of $d(l, p_s) < d(l, p_i)$ and $\sum_{j=1}^{n}(T_{sj} \to D_s) > \sum_{j=1}^{n}(T_{sj} \to D_i)$, as mentioned in section 1.2. Then based on the basic idea, we have the following definitions.

Definition 3 (Equidistant uncertainty position). For a location l and a set of PoIs $P = \{p_1, p_2, ..., p_n\}$, if the distance of location l to random k points in this set satisfies $d(l, p_1) = d(l, p_2) = ... = d(l, p_k)$, then these points in this set are called equidistant positions. Furthermore, if the distance value $d(\cdot)$ in the distance range is involved in $[1, \max]$ and uncertain, these points are called the equidistant uncertainty position.

Definition 4 (Multi-destination navigation). For at least n segments in a trajectory of a route, if in the navigation each segment correlates to a different destination, i.e. $T_{s1} \to D_1 = T_{s2} \to D_2 = ... = T_{sn} \to D_n$, then this navigation can be called multi-destination navigation.

Based on principles of definition 3 and definition 4, in this paper, we choose concentric circles and the Voronoi diagram to realize equidistant uncertainty position as well as multi-destination navigation, so as to provide personal privacy preservation service for the user in ORHN.

Definition 5 (Concentric circles). We call the circles with the same center but various radiuses as the concentric circles. In each annulus there will distribute a number of PoIs with a certain radius, i.e. in the distance interval $[1, d]$ all PoIs are distributed on various annuluses in the range of $\int_1^d 2\pi r d_r$.

Definition 6 (Voronoi diagram). Assume that a graph structure $G = (V, E, C)$ is constructed by a group of lines between points of every two adjacencies, where V denotes the vertexes and E denotes the edges of each grid, and then C denotes the center points of each grid. In addition, as each grid covers a number of PoIs, the distance between every PoI to the center point in the same grid is shorter than PoIs to center points of other grids, i.e. the distance satisfies $d(E_{current}, C_{current}) < d(E_r, C_{current})$ and $d(E_{current}, C_{current}) < d(E_{current}, C_r)$. Where $E_{current}$ and $C_{current}$ denote the edge and center point in the current grid, and E_r and C_r denote random edge and the center point of other grids in this Voronoi.

Then the IRDP Ride system for ORHN is designed based on the two definitions mentioned above.

III. The IRDP Ride system

A. The whole process of the IRDP Ride system

If a user utilizes the IRDP Ride system, he/she just chooses the initial point and the destination in the map, and then set several privacy parameters in this system. The system calculates the secure pick-up and drop-off points according to the privacy parameters and then sends the visual result to the user. After that, the user selects the points in the annulus and adjusts the location of some crucial grids in Voronoi, and feeds back these chosen to this system. Then the IRDP Ride system selects an unoccupied taxi in the adjacent range of the pick-up point and sends the ride information to both the user and the taxi driver. At the same time, the IRDP Ride system also sends multi-destination navigation information to the taxi driver, when the taxi moves into a new grid or leaves the

center point of the current grid until the user arrives at the drop-off point. If a sharer wants to participate with the user, and the user agrees to share with this sharer, the IRDP Ride system calculates whether their destination is adjacent. If so, this system sends another Voronoi to the sharer and repeats the calculation of secure pick-up and drop-off points and multi-destination navigation, and then navigates the taxi with different grids in two Voronoi. The whole process of the IRDP Ride system can be depicted in Fig.2, and the detailed operation will be elaborated in the following section with three algorithms.

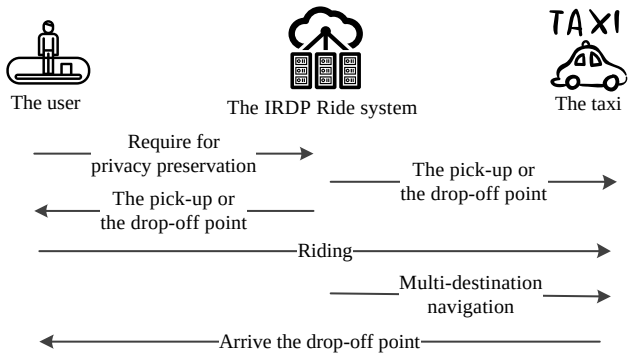


Fig. 2.The process of the IRDP Ride system.

In addition, as the IRDP Ride system is mainly focused on privacy violations from the taxi driver and the sharer, the security of information transmission among three entities is not considered, so we assume there is a secure channel that can be used to transmit information.

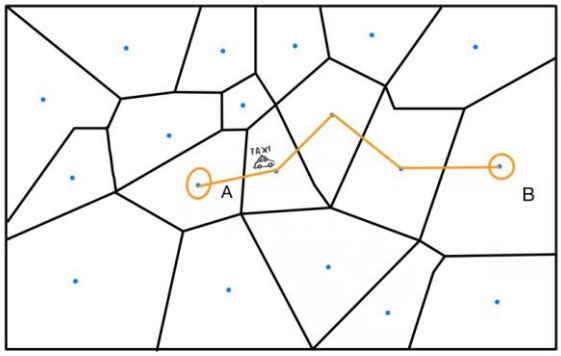


Fig. 3.The result of the IRDP Ride system.

Combine the function of equidistant uncertainty position and the function of multi-destination navigation in the IRDP Ride system. In this section, an example based on figure 3 is given to further elaborate the process. As shown in figure 3, suppose that a user is located in a random location at an uncertain annulus around the center point of A, and the destination of this user is in a random location at an uncertain annulus around the center point of B. If this user wants to preserve his/her privacy and utilizes the IRDP Ride system, he/she will get the secure pick-up and the secure drop-off point shown as the center point of A and B. At the same time, he/she will also get the navigation range of A and B partitioned with a Voronoi diagram and the routing is the broken line from A to B. As the points of initial and destination are generalized by two groups of concentric circles, the adversary cannot confirm where the PoIs located, the range attack will fail and the privacy of these two PoIs will be preserved. Then as different center points in various grids, the navigation from the edge to this point or from this point to the other edge will direct to different PoIs, the segments direction attack will fail and the privacy of the routing will be preserved.

B. The algorithm for preserving the initial and destination point

In the IRDP Ride system, the primary function is equidistant uncertainty position, i.e. secure pick-up and drop-off points find. Then based on this idea, in this paper, we utilize concentric circles to realize the uncertainty. In Algorithm 1, we elaborate on the process of finding the equidistant uncertainty position with the input of the initial point or destination as well as several privacy parameters.

Algorithm 1 The algorithm of concentric circles (calculating the secure pick-up or drop-off point).

Input: The initial point or the destination l, the maximum moving distance $d_{\max}$, the anonymous value k (the number of PoIs in the uncertain annulus)

Output: The pick-up point l_i or the drop-off point l_d

1. Creates an annulus A_1 with the point l as the center point, the maximum moving distance $d_{\max}$ as the radius;
2. Set=All PoIs on this A_1;
3. **While** (Set !=Null) **Do**
4. Selects a random point p_r in annulus A_1, and then creates another annulus A_2 that passes the initial point with this point as the center point and $d_{\max}$ as the radius;
5. **While**(the number of PoIs on A_2 < k) **Do**
6. $r = d_{\max} - 1$; //Decrease the value of radius.
7. A_2 =a new annulus with r as the radius and p_r as the center point;
8. **If** (r<=1)
9. Return (This point is failed to generalize the real PoI.);
10. **End if**
11. **End while**
12. Set=Set- p_r ;
13. **End while**
14. **If**(Set==Null and all PoIs do not satisfy the requirement of generalization)
15. Return (The system failed to provide privacy preservation service.);
16. **Else**
17. Return (p_r); // The point p_r will be the secure pick-up point l_i or the drop-off point l_d .
18. **End if**

As depicted in the process of Algorithm 1, the user will get the secure pick-up and drop-off points, i.e. the center point of

several groups of concentric circles as shown in figure 3. With the center point as the pick-up or the drop-off point, the real PoIs will be generalized by all PoIs in the circle of the largest annulus, and the adversary will be difficult to identify the real PoI in the generalized set. The security of the concentric circles will be elaborated on in the section on security analysis.

C. The algorithm for Segments navigation

Another primary function of the IRDP Ride system is multi-destination navigation, i.e. making various segments point to different destinations. Based on this idea, we utilize the Voronoi partition and the navigation shifting to complete multi-destination navigation. In this section, we divide this function into two different stages: the stage of Voronoi partition and the stage of multi-destination navigation. In the stage of Voronoi partition, the IRDP Ride system first confirms the range of the navigation area with the pick-up and drop-off points, and then the IRDP Ride system expands the navigation area and divides it into several groups of grids. Finally, the IRDP Ride system calculates the shifted navigation for each middle grid. We show the elaborated process of Voronoi partition in Algorithm 2.

Algorithm 2 The algorithm for partitioning the navigation area.

Input: The secure pick-up point l_i or the drop-off point l_d, the range R_1 that constructed by l_i and l_d, the anonymous value of k_2 and k_3 (the anonymous value for grids in Voronoi and grids between the pick-up and drop-off points).

Output: The expanded area R_2 with the partitions by Voronoi grids.

1. Expanding the area R_1 into a larger range R_2 with random areas;
2. Num=0; //Used to store the number of grids between l_i and l_d temporarily;
3. **While** (Num< k_3) **Do**
4. Randomly selects k_2-2 points in R_2, and creates the Voronoi diagram in this region with the random points and the points of l_i and l_d;
5. Num=the number of point between l_i and l_d;
6. **End while**
7. Return(R_2 with the partitions by Voronoi grids);

With the process of Algorithm 2, we will get the navigation area and its Voronoi partition. In the stage of multi-destination navigation, the IRDP Ride system will calculate the shifted routing for all middle grids, and send the navigation result to the taxi driver once the car moves into a new grid. The detailed process of shifted navigation is shown in Algorithm 3.

Algorithm 3 The algorithm of multi-destination navigation.

Input: R_2 with the partitions by Voronoi grids, the secure pick-up point l_i or the drop-off point l_d denoted as p_r.

Output: The navigation destination D_c for current grids.

1. Flag= k_3;
2. **While** (Flag!=0) **Do**
3. D_c =the center point of the next Voronoi grid;
4. Gives the path from p_r to D_c;
5. $p_r = D_c$;
6. Flag = Flag -1;
7. **End while**

The IRDP Ride system will send the navigation result to the taxi driver until the car arrives at the real drop-off point. Before the user gets off, no one knows whether the current navigation points to the real drop-off point, and each section of the shifted navigation may point to a different destination. As a result, the route and destination of the user will be preserved.

D. The algorithm for resisting attack from the sharer in the same car

In Algorithms 1-3, two stages of the IRDP Ride system are elaborated and we consider that these algorithms are designed to cope with attacks from the taxi driver and the sharer. However, whether the sharer can share the taxi with the user is determined by the uniformity of their route and destination. The IRDP Ride system has to give judgment to car sharer, if the user wants to share with the sharer. In addition, the IRDP Ride system also has to confirm the privacy security for both the user and the sharer. The detailed process of sharer judgment is shown in Algorithm 4.

Algorithm 4 The algorithm of car sharer judgment.

Input: The destination D_c of current grid and the drop-off point l_d of the current user, the pick-up point l_i' and the drop-off point l_d' of the sharer, the expanded area R_2 with the partitions by Voronoi grids of the current as well as the expanded area R_2' with the partitions by Voronoi grids of the sharer.

Output: Whether this taxi can be shared

1. The current grid G_1 and the pick-up grid of the sharer G_1', the drop-off grid G_2 of the user and the drop-off grid G_2' of the sharer.
2. **If(** $D_c \in G_1'$ and $l_i' \in G_1$ and $l_d \in G_2'$ and $l_d' \in G_2$)
3. Return(this car can be shared);
4. **Else**
5. $G_1 = G_1$ + grids in vicinity;
6. $G_2 = G_2$ + grids in vicinity;
7. $G_1' = G_1'$ + grids in vicinity;
8. $G_2' = G_2'$ + grids in vicinity;
9. **End if**
10. **If(** $D_c \in G_1'$ and $l_i' \in G_1$ and $l_d \in G_2'$ and $l_d' \in G_2$)
11. Return(this car can be shared);
12. **Else**

13. Return(this car cannot be shared);
14. **End if**

If the result of Algorithm 4 is this car can be shared and the user wants to share with the sharer, the IRDP Ride system adds another route segment from the current location D_c to the pick-up point l_i' of the sharer. Then when the taxi arrives at the drop-off point l_d and the user gets off this taxi, the IRDP Ride system also adds new navigation to the drop-off point l_d' of the sharer. Other dispositions for the sharer are similar to the user as shown in Algorithm 1-3.

IV. THE SECURITY AND STABILITY ANALYSIS

A. The resistance to range attack

In this paper, we consider both the taxi driver and sharer two entities as adversaries, but for different adversaries, the background knowledge that they grasp is different and they can utilize different attacks to infer the privacy of the user (such as the range attack and segments direction attack). Then, in this section, based on the variety of background knowledge (such as the knowledge of pick-up and drop-off points, the knowledge of algorithms in this system, and the knowledge of parameters used in this system), we first discuss the resistance of the IRDP Ride system for range attack.

1) The resistance to the knowledge of the pick-up and drop-off points

We first assume that the adversary has got the pick-up and drop-off points of the user. A great majority of adversaries will have this type of background knowledge, and they will know where the user gets on or off. In this condition, the adversary will consider these points as the initial point and destination or utilize these points to infer the initial point or destination with the range attack (Definition 1). For the pick-up point l_i, the adversary can calculate $d(l_i, p_s) < d(l_i, p_i)$ to guess the nearest location as the real PoI. However, in the IRDP Ride system, as disposed of by algorithm 1, the user has moved an uncertain length of distance d. As the real PoI is generalized by at least k PoIs on the real annulus and several groups of annuluses in or out of this annulus, it will be more difficult for the adversary to identify the real PoI. Suppose that, if there are n annuluses and k PoIs in each annulus, the number of PoIs that the adversary gets is $nk\int_1^d 2\pi r d_r$, as the adversary cannot obtain the distance, we consider that $d \to \infty$, there will be infinite PoIs that the adversary cannot distinguish. Thus, with the pick-up and drop-off points as background knowledge, the adversary cannot identify the real PoI and the IRDP Ride system can resist the range attack.

2) The resistance to the knowledge of algorithms in this system

As all algorithms in the IRDP Ride system are opened to all users, the adversary will also get the process of algorithm 1 and infer the length of moving distance with the ability of the user. However, in algorithm 1, the real PoI must not be located in the outermost annulus of concentric circles, so the radius of the real annulus will be a random number in the range of $r \in [1, d_{\max}]$, and the number of PoIs that the adversary gets is also $nk\int_1^{d_{\max}} 2\pi r d_r$, where $d_{\max}$ is the longest distance that the user can move. As a result, the adversary cannot identify the real PoI and the IRDP Ride system can resist the range attack with this type of background knowledge.

3) The resistance to the knowledge of parameters used in this system

We also assume that the adversary knows the precise value of each parameter, he can get the real distance as well as the real anonymity value of k, but he is still unable to confirm the precise PoI. In this condition, the adversary will know which is the real annulus and get the number of PoIs in the range of $[k, 2\pi d]$, but in $2\pi d$ points which k are selected and whether the real PoI is in these k points is still uncertain. So the probability of identifying the real PoI is in the range of $[1/k, 1/2\pi r]$, even if $k = 2$, the probability of successfully identifying the real PoI is still not higher than the probability of the random coin toss.

In conclusion, based on the resistance of the IRDP Ride system for range attack, even if the adversary has gotten above three types of background knowledge, he still has a lower probability to identify the real PoI, so the IRDP Ride system can provide privacy preservation service for the ORHN user.

B. The resistance to segments direction attack

When moving to the drop-off point, the taxi driver and the sharer can obtain the unified direction of various segments and the predicted drop-off point with sub-segments previously as background knowledge. Then with these two types of background knowledge, the adversary can utilize the segments direction attack to infer the privacy of the user. Therefore in this section, we will discuss the resistance of the IRDP Ride system for segment direction attacks with two types of background knowledge mentioned above.

1) The resistance to the knowledge of the unified direction of various segments

In the whole trajectory to the destination, the segment can be divided into small sub-segments and denoted as $T = \{T_{s1}, T_{s2}, ..., T_{sn}\}$. Then for a specified point D_s, if each sub-segment satisfied $\sum_{j=1}^{n}(T_{sj} \to D_s) > \sum_{j=1}^{n}(T_{sj} \to D_i)$, the point D_s can be considered as the destination and the trajectory of the route can be predicted in advance. However, in the IRDP Ride system, with the help of multi-destination navigation, each sub-segment will be navigated to a different center point of the Voronoi grid, and then for the set of destination $D = \{D_1, D_2, ..., D_k\}$, the correlation of sub-segments and destination will satisfy $(T_{s1} \to D_{s1}) = (T_{s2} \to D_{s2}) = ... = (T_{sk_2-1} \to D_{sk_2-1})$, i.e. in

the whole trajectory there will be at least k_2-1 sub-segments direct to k_2-1 PoIs and the probability are equal to each other. As a result, only if the user gets off this taxi, the adversary will not get the real drop-off point during the whole journey. Furthermore, even if the adversary obtains the drop-off point, it will also be difficult to identify the real PoI of the user that this user from or wants to visit. So the IRDP Ride system can resist segments direction attack under the background knowledge of the unify direction of various segments.

2) The resistance to the knowledge of the prediction for potential sub segment

Another type of background knowledge for segments direction attack is the predicted drop-off point with sub-segments previously. Suppose that the PoIs that the taxi has passed can be denoted as $P=\{p_1, p_2, ..., p_i\}$, and the next set of PoIs can be denoted as $P=\{p_1^{'}, p_2^{'}, ..., p_n^{'}\}$, and then for at least n potential PoIs p_m the real point may satisfy $p_{i+1}=\max(pro(p_1, p_2, ..., p_i \mid p_m))$, where $\max(\cdot)$ is the maximum value of probability. However, in the IRDP Ride system, as there is no correlation between the previous PoIs and the next point, the correlation probabilities of these PoIs will be equal to each other, i.e. $pro(p_1, p_2, ..., p_i \mid p_1^{'}) = ... = pro(p_1, p_2, ..., p_i \mid p_n^{'})$. As a result, the adversary will be difficult to infer the real point next, and the IRDP Ride system can resist segment direction attack under the background knowledge of the prediction for a potential sub-segment.

C. The stability analysis

In this paper, we consider the stability of the IRDP Ride system on both the distance of the user's moving and the distance of route shifting, as the former is determined by the mobility of the user, and the latter is determined by the time consumption of ORHN.

1) The distance of the user's moving

For the distance of the user's moving, as the pick-up point and the drop-off point are used to generalize the initial point and the destination, the user has to move to the specified location with the maximum distance $d_{\max}$. So during the whole process of ORHN, the distance of a user's moving is in the range of $(0, 2d_{\max}]$, i.e. in the worst scenario, the user has to move the length of $2d_{\max}$. However, as the length of moving is determined by the user, if the user wants less level of privacy preservation and better service, he/she can move a shorter distance. So the distance of the user's moving cannot affect the stability of the IRDP Ride system. In addition, as the PoIs used in generalization may not be the outermost annulus, the user may not move to the maximum distance to generalize the initial point or destination.

2) The distance of route shifting

During the journey of ORHN, the taxi will be guided to a group of different PoIs, so as to resist segments direction attack. As a result, the taxi has to move more miles to reach the destination. However, in the IRDP Ride system, PoIs used in multi-destination navigation are randomly selected by the grids of the Voronoi diagram, which will limit the length of excess distance and reduce the influence of shifting on the stability of this system. In a grid of Voronoi, the distance of PoIs in this grid to the center point is the shortest to these PoIs to the center point of other grid, so in this grid, the distance of shifting is shorter than the distance of shifting to other points. Then in the whole process of ORHN, the total length of shifted distance is not much longer than the real trajectory. So the distance of shifting does not affect the stability of the IRDP Ride system.

V. Performance Analysis

A. Preparation

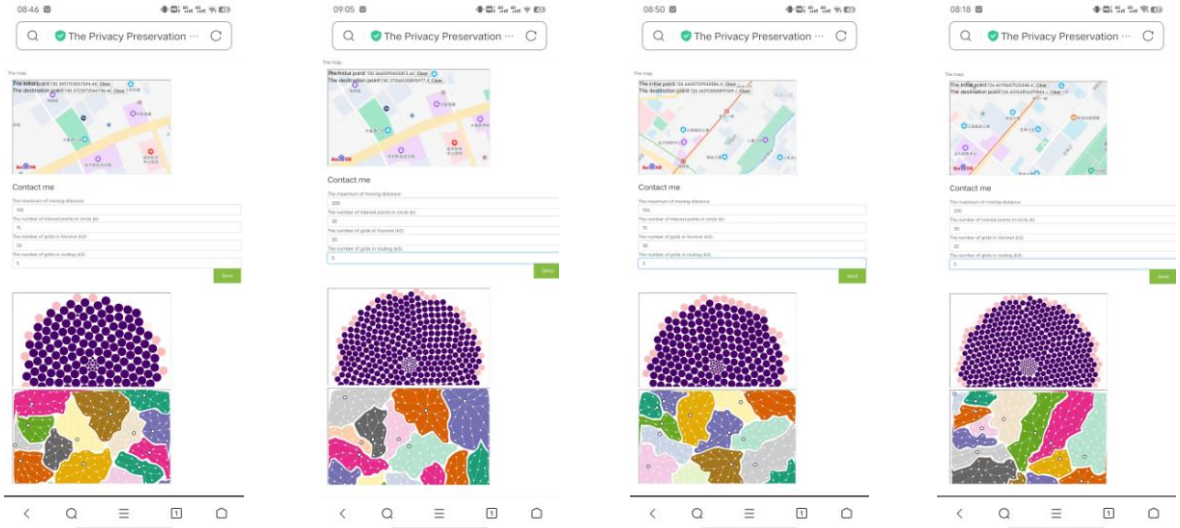
As the IRDP Ride system is designed for providing privacy preservation service when the user utilizes ORHN, in this section, we will infer the performance of this system with several groups of experiments in both application and comparison. We first deployed this system in two different types of cities, such as the city of Harbin and the city of Jiamusi (a provincial capital and a general city). Then we discuss the difference in parameters used in various cities and discuss the optimal parameters in different cities. Finally, with the results of the comparison, we discuss the superiority of the IRDP Ride system. All experiments are implemented by a laptop and a mobile phone. The configuration of the laptop is CORE i7 1.8GHz, 8GB of memory, and the operating system is Microsoft Windows 10. The configuration of the mobile phone is 2.6GHz, 12 GB memory, and Android 11 system, which is used to get the application result. In the application of the IRDP Ride system, we utilize JavaScript as the development language for web and the Python 3.6 as the development language for functions. For the API, we utilize Baidu map and Observable, and we calculate the distance of moving and navigating with these two APIs. The PoIs of Harbin and Jiamusi are collected from Baidu, and they are used to infer the optimal parameters for the distance of the user's moving, the breadth of the annulus, the distance of taxi shifting as well as the number of PoIs, etc. Then the result will show the different values of parameters in different cities and the optimal parameters used in different environments (such as the large city and the small city).

The comparison is focused on the performance of protected locations, the running time as well as the success ratio of service providing. Other systems used for comparison are the destination shifting ε-sensitive indistinguishable system [16], the taxi share PGRide system [19], the trajectory shifting TrackU system [23], the prediction matching pRide system [24], the personalized taxi enhanced Private Ride system [18], the navigation generalization NLPPMA system [25] as well as the IRDP Ride system. Then the comparison result will show the superiority of the IRDP Ride system.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2022.3232504

B. *Applications*

For testing the parameters used in the IRDP Ride system, we first apply it in the city of Jiamusi. In this application, we set the maximum distance of moving as 100 meters, and the values of private parameters are k=15, $k2$=20 and $k3$=5. The comparison result is applied with the maximum distance of moving as 200 meters, and the values of private parameters are k=30, $k2$=20 and $k3$=5. For another group of comparison, we apply the IRDP Ride system in the ORHN in the city of Harbin, and the values of private parameters are set as we have used in the city of Jiamusi. Then a group of results is shown in figure 4. In this figure, the part of location selection is the result of the Baidu map, which is used to choose the initial point and the destination. A group of input boxes is used for setting the private parameters (such as the maximum distance, k, $k2$, $k3$). The one of other two parts is used to show the number of PoIs in the circle and on the annulus (the pink points are the PoIs located on the annulus), and the other is used to show the number of grids in the Voronoi diagram and the number of grids between the pick-up and drop-off points (the grids with the same color are the grids with the pick-up and drop-off points, and the white points are PoIs in this grid). Then we show the comparison result in subfigures 4a, 4b, 4c and 4d to demonstrate the difference of utilizing the IRDP Ride system in two different cities with different parameters. In addition, in the subsequent parts of this paper, the results of influence for testing these parameters are also calculated in the above environments with similar settings.



a. The result in Jiamusi with 100 meters.
b. The result in Jiamusi with 200 meters.
c. The result in Harbin with 100 meters.
d. The result in Harbin with 200 meters.

Fig.4. The result of application in two different cities.

Fig.5 shows the influence of moving distance (the maximum radius) on the number of PoIs in two cities. From this figure, we can see that the long distance moving the more PoIs in this circle, and if the moving distance is longer than 10 meters, the number of PoIs included in this circle will increase dramatically, and this number will be steadily increasing once the moving distance reaches a certain distance (such as 200 meters in Jiamusi and 500 meters in Harbin). In addition, with the same length of moving distance in different cities, the number of PoIs included in the circle is different from each other, just as the difference of curves shown in this figure. The reason for the differences mentioned above can be briefly concluded as the total number of PoIs and the distance between PoIs in different cities. As a long distance of the user's moving will bring in a larger range of circular areas and then more PoIs will be included in this circle, the number of PoIs will be increasing along with the increasing moving distance. In addition, as the PoIs also take up a certain amount of space, the increasing trend will be increasing dramatically or steadily with the range of circles and the space of PoIs take. Another phenomenon can be concluded as the following reason. We know that Harbin is a provincial capital and compared to a general city (such as Jiamusi)it is much larger, and the space for a PoI is also larger than similar PoI in Jiamusi, so in the same range of circle less PoIs are included in this circle.

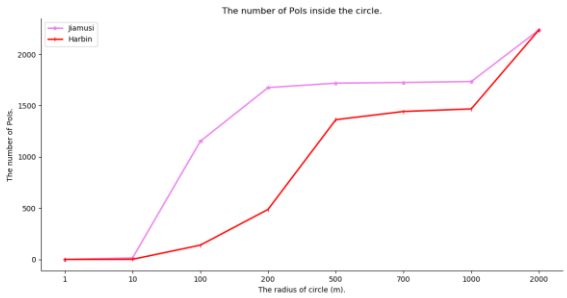


Fig. 5.The length of radius vs. the number of PoIs.

Fig.6 shows the influence of annulus width (the maximum width) on the number of PoIs in two cities. From this figure, we can see that the number of PoIs is increasing with the width of the annulus, and the number of PoIs is still lower in Harbin than in Jiamusi with the same width of the annulus. In addition, if the width of the annulus is higher than 100 meters, there will be more than 500 PoIs that can be used to generalize the real point. If the annulus width is higher than 100 meters, a bigger circle with an even higher length of moving distance is needed, so it is better to choose this parameter as a value between 100 to 1000 meters as shown in this figure. The

reason for these results can be concluded as the wider annulus will bring in more PoIs in its range, and the distance between PoIs and the space of PoIs is larger in the provincial capital than in a general city.

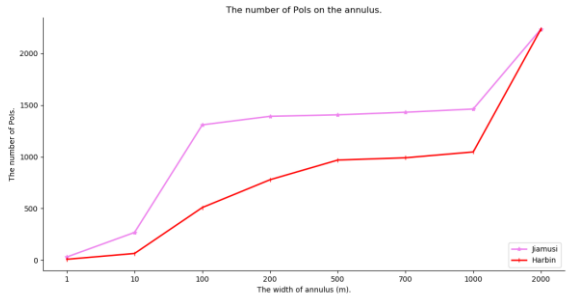


Fig.6. The width of annulus vs. the number of PoIs.

Fig.7 shows the value of correlation that calculated by the increased number of PoIs correlates to the drop-off point, and this value is used to measure the degree to of an adversary identifies the real destination. In this paper, entropy is used as the metric to measure this value and the value is calculated by $H(i) = -\sum_{i=1}^{k_3} p(i)\log p(i)$, where $p(i)$ denotes the probability of correlating the drop-off point with the increasing number of PoIs. Then the higher value of $H(i)$ indicates the more difficult the adversary identifies the drop-off point. In this figure, the Baseline indicates a route without privacy preservation, and the Jiamusi and Harbin indicate two routings in the cities of Jiamusi and Harbin with the help of the IRDP Ride system. As the increasing number of PoIs cannot be used to improve the probability of correlating the drop-off point so the adversary can just estimate the drop-off point by the number of grids in the vicinity of the current grid, the value of $H(i)$ cannot change with the increasing number of PoIs between the pick-up and the drop-off points.

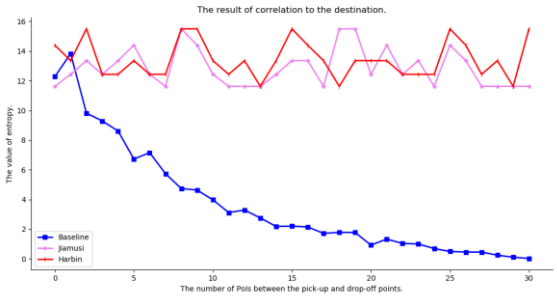
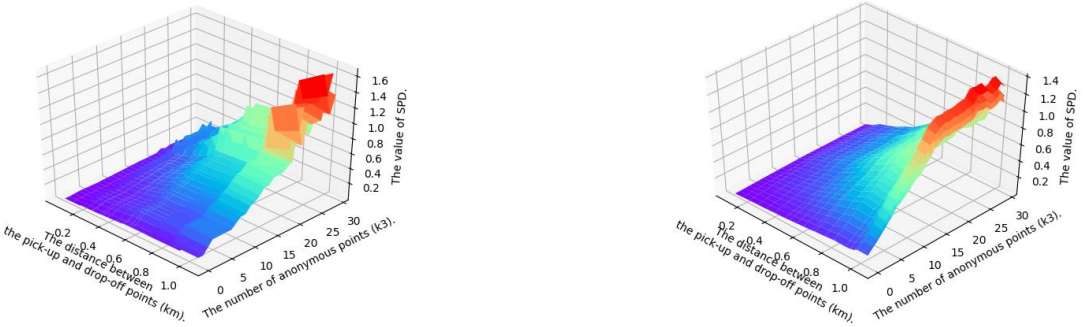


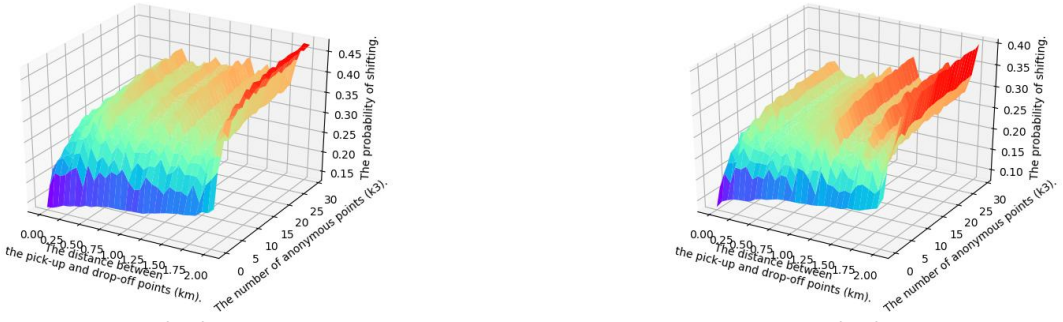
Fig.7. The value of entropy vs. the number of PoIs.



a. The value of SPD in Jiamusi.

b. The value of SPD in Harbin

Fig.8. The value of SPD vs. the number of anonymous points *k*3 and the distance between the pick-up and drop-off points.



a. The probability of shifting in Jiamusi

b. The probability of shifting in Harbin

Fig.9. The probability of shifting vs. the number of anonymous points k3 and the distance between the pick-up and drop-off points.

Fig.8 shows the value of SPD. This value is used to measure the similarity of the trajectory to the shortest route between the pick-up and drop-off points, and this value is called the sum of pairs distance (short for SPD) and is calculated by $SPD(A,B) = \sum_{i=1}^{k_3} d(a_i, b_i)$. From the

comparison of the two sub-figures in this figure, the value of SPD is increasing with the increase of *k*3 and the distance between the pick-up and drop-off points, which means a lower similarity to the shortest route. Then, as the interval of PoIs in Harbin (Fig.8 b) is shorter than in Jiamusi (Fig.8 a) (more PoIs in Harbin than Jiamusi), the total value of SPD in Jiamusi is higher than in Harbin, which means a lower similarity for the trajectory generated in Jiamusi than Harbin.

Fig.9 shows the influence of the length of the total distance between the initial point and the destination on the probability of shifting. This value is calculated by $P_{shift} = (dis(i,d) - dis(i',d'))/dis(i,d)$, where $dis(i,d)$ is the distance between the initial point and the destination, and $dis(i',d')$ denotes the distance between the pick-up and drop-off points. As this value is influenced by the parameters of *k*3 (the number of grids between the initial point and the destination) and the summation of the distance between the pick-up and drop-off points, we utilize the three-dimensional diagram to demonstrate the changes. From this figure, we can see that for different values of *k*3, the more the number of grids between the initial point and the destination the higher value of the probability of shifting, as more grids bring in a longer length of moving distance. Then, we can also see that along with the increase in the distance between the pick-up and drop-off points the value of the probability of shifting is increasing significantly, as the shifting will bring in a longer journey than straight movement. In addition, compared with the two sub-figures, we can see that the probability of shifting in Jiamusi (Fig.9 a) is higher than in Harbin (Fig.9 b), as the interval of PoIs in Harbin is shorter than in Jiamusi and the probability of shifting in Harbin is also lower than in Jiamusi, as the distribution of road network in provincial capital is much more complete than the road network in a general city.

TABLE I
THE COMPARISON OF DIFFERENT SYSTEMS IN PROTECTION FIELDS

Various systems	The main strategy	The protections for privacy fields		
		The initial point	The trajectory	The destination
ε-sensitive indistinguishable system [16]	Differential noises	×	√	√
PGRide system [19]	Homomorphic encryption	√	×	×
TrackU system [23]	Location shifting	√	×	√
pRide system [24]	Homomorphic encryption	√	×	×
enhanced Private Ride system [18]	Range obfuscation	√	×	√
NLPPMA system [25]	Shifting and generalization	√	×	√
IRDP Ride system	Obfuscation and shifting	√	√	√

In order to compare the performance of the IRDP Ride system with others, we uniform the parameters for different strategies and then get the result of running time and the result of success ratio. For these parameters, we utilize the same value for the number of generalizing users and noises, the same value for the digits of the encryption key, the same value for the distance of shifting as well as the same value for the range of obfuscation areas.

Fig.10 shows the result of running time compared with other systems. From this figure, we can see that the running time for systems with the strategy of encryption is higher than others (such as the PGRide system, and the pRide system), and the degree of increase is steadily along with the increase of the value of the private parameter. The running time for systems with the strategy of noise addition, generalization, and obfuscation is a bit lower than encryption (such as the ε-sensitive indistinguishable system, the enhanced Private Ride system, and the NLPPMA system), and the degree also increases more dramatically than encryption. Then running time for systems with the strategy of shifting is the lowest (such as the IRDP Ride system and the TrackU system), and the trend of increasing with the increasing value of the private parameter is also steady. The phenomenon mentioned above can be summarized for the following reasons. For systems with encryption, as the process of encryption is more complicated than other strategies (such as noise addition, generalization, obfuscation as well as shifting), the running time is higher than these systems. For systems with shifting, as just the user's location or route is changed which is less time used in private operation than the noise addition, generalization as well as obfuscation, the running time is lower than these systems. In addition, as there is less complexity change with the increasing of the digits of the encryption key and the increasing of shifted distance, the trend of running time increasing with the increasing value of the private parameter is lower than systems with noise addition, generalization as well as obfuscation.

Fig.11 shows the result of the success ratio for these systems. The value of the success ratio is calculated by the proportion of successfully providing privacy preservation services for users. From this figure, we can see that the success ratio of systems with the strategy of encryption is the highest, as these systems do not need to find other users or add noises (such as the PGRide system, or the pRide system). The

C. Comparison results

In order to facilitate the comparison with other similar systems, we show the comparison result in Table 1. In this table, we make the comparison in the following aspects, such as the main strategy, the protections for privacy fields (the initial point, the destination as well as the trajectory), and so on. From this table, we can see that just the IRDP Ride system can provide privacy preservation service for every field during the whole journey of ORHN, and other systems usually just consider a limited number of these fields. Then from the comparison of the main strategy that these systems used, we can see that the IRDP Ride system is based on obfuscation and shifting, which shows a better result on time efficiency than systems with encryption and noise addition (this result will be further demonstrated in the following paragraph). So we can consider that the IRDP Ride system will be easier to be implemented and deploy in the real environment.

success ratio of systems with shifting is a bit lower than systems with encryption (such as the IRDP Ride system and the TrackU system), as these systems have to find the available locations or routes for user's shifting. The success ratio of systems with the strategy of noise addition, generalization, and obfuscation is the lowest (such as the ε-sensitive indistinguishable system, the enhanced Private Ride system, and the NLPPMA system), as sometimes these systems fail to add noises or find anonymous users or extend the available region.

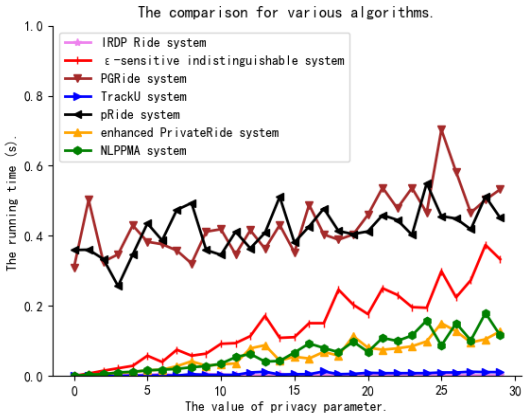


Fig.10.The comparison of different systems in running time

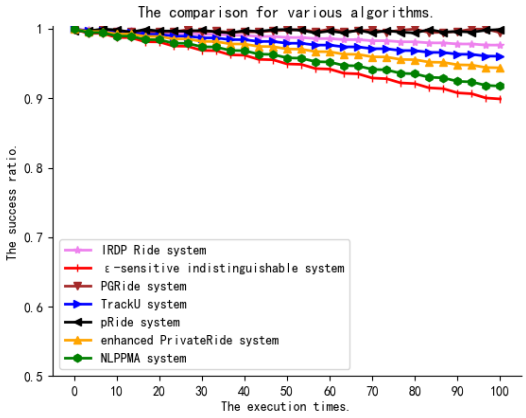


Fig.11.The comparison of different systems in success ratio

In conclusion, compared with several systems used currently, we can consider that the IRDP Ride system performs better than these systems in both the running time and the success ratio. Furthermore, the IRDP Ride system can also provide the service of privacy preservation for users in more fields (such as the initial point, the destination as well as the trajectory).

VI. Related Work

As we have mentioned in the section introduction, the algorithms and systems used in privacy preservation can be briefly classified into two categories: privacy preservation for requirement and privacy preservation for service. In this section, we will briefly introduce several related works in each category and discuss the advantages and disadvantages of each system, and then a brief analysis of these systems is given.

A. Privacy preservation for requirement

The service of privacy preservation for a requirement can be seen as a type of passive service, i.e. the service provider just provides privacy preservation service once the user requires LBS. Under this definition, the algorithms and systems in this category can also be classified into the type of snapshot query and the type of continuous query. In the type of snapshot query, Sun et al. [26] proposed a users collaboration method with the location label, so as to solve the problem of finding collaborative users with similar attributes in the vicinity. Hua et al. [27] utilized the Geo-indistinguishable to perturb the probability of frequent queries in the same location. Based on the conception of dividing sensitivities, Yin et al.[28] proposed a privacy protection method to equilibrate the quality of service and privacy protection. Li et al. [29] utilized the range of two-dimensional maps to achieve location privacy, query privacy as well as semantic privacy. Guo et al. [30] formulated the problem of encrypted geographic queries as range-based pattern matching and carefully craft security schemes to enable efficient range queries in the ciphertext domain. Han et al. [31] utilized the method of attribute-based encryption to share personal location with others in the condition of adaptive privacy preservation. Zhang et al. [4] utilized the distribution of blockchain to collect collaborative users in different areas to generalize the real location. However, along with the development of LBS, more and more people prefer to require continuous results along the journey, so in the category of privacy preservation for requirements, more algorithms and systems were focused on the privacy preservation service in continuous query.

In the type of algorithms or systems for continuous query, Ye et al. [32] considered that the privacy preservation system must provide diversity in the continuous query along the user's journey, so as to ensure the precise query can be effectively generalized. Based on the result of the prediction, Sepahkar et al. [33] proposed a generalization model for collaborative users during the whole journey of continuous queries. Zhang et al. [34] provided a cache model to provide continuous generalization with *k*-anonymity, so as to reduce interaction with the un-trusted service provider. Kalaiarasy et al. [35] utilized the mechanism of variant ring signature to change the pseudonym of vehicles in mixed zones of VANET to preserve personal privacy in continuous queries. Based on the conception of differential privacy and non-uniformity, Deldar et al. [36] proposed an algorithm for location un-linkable to prevent the discrete locations from being constructed into trajectory. Cao et al. [37] considered that the generalization in continuous query had to pay attention to the Spatio-temporal event in the user's journey, so as to prevent the adversary from identifying the real trajectory.

Although the algorithms and systems in the category of privacy preservation for a requirement can provide better service for preserving user's privacy in the snapshot query and continuous queries, in other types of service, such as in the service without query, there is no query sent to the service provider and no collaborative users or queries can be used in

generalization, these algorithms and systems seem a bit out of their scope. As a result, algorithms and systems in the category of privacy preservation for service are prosperous and we will discuss their advantages and disadvantages of them in the following sub-section.

B. Privacy preservation for service

As the service is provided by the service provider without any query, this type of service can be seen as a type of active service, and then based on the differential of applications it can also be considered as the service on road networks and the service on other applications. In the type of service on road networks, the algorithms and systems of privacy preservation mainly focus on privacy violation in the journey of users moving. In the road networks, Wang et al. [38] proposed a trigger-based pseudonym exchange scheme, which to prevent this user from being tracked in network monitoring. Based on the scheme of encryption and cloud computation, Yang et al. [39] provided a system for privacy preservation k nearest neighbor finding on road networks. Li et al. [40] utilized the differential privacy model to provide a pseudonym swap algorithm, which can reduce the probability of linking the same vehicle in VANETS. Khodaei et al. [41]considered that just the simple mix zones in road networks are not enough, and more users with similar attributes are needed to enhance the collaboration of generalization and enhance un-traceability. Ullah et al. [5] considered that the distributed mix-contexts in road networks will conceal the real intention of the user, and proposed a system to provide context generalization.

The algorithms and systems mentioned in the previous paragraph can be seen as privacy preservation services for users in the application of road networks, in other types of applications more algorithms and systems are proposed (such as the application of mobile phones, the application of health care and the application of crowdsensing, etc), and they can be seen as important supplements for privacy preservation service. In the application of mobile phones, Zhang et al.[42] proposed a path-shifted system for map services on mobile phones, so as to prevent the map provider from identifying the real intention of navigation in map service. Natgunanathan et al. [43] proposed a privacy protection smart health care system, which can provide health care service without location privacy violation. Based on the conception of generating decoy queries to hide the real users' locations, Kang et al.[25] proposed a novel location privacy preservation mobile app to protect mobile phone users. In the application of crowdsensing, Zou et al. [44] proposed a blockchain-based crowdsensing system to preserve the privacy of mobile phone users. Wang et al. [45] proposed a system with a truthful incentive mechanism, so as to effectively distribute the crowdsensing tasks. Zhao et al. [46, 47] proposed privacy preservation schemes with quality-aware and reliability-aware to further improve the incentive in crowdsensing tasks. Sadhu et al. [48] proposed a multi-modal collaborative mobile phone privacy preservation algorithm, which can cope with several attacks that are used to violate mobile phone users' privacy. In the application of health care, in order to cope with the privacy violation in Covid-19 contact tracing, Bhardwaj et al. [13] proposed a system to find the contact user without violating location privacy on mobile phones.

Besides these applications, privacy violations also threaten online ride-hailing services, and some algorithms or systems became concerned about this issue. As we have mentioned in the section introduction, some algorithms or systems that combined privacy preservation in the process of sharing and tracing have been proposed. Luo et al. [20] utilized the cryptographic primitives to securely and efficiently estimate the shortest distances between riders and drivers in road networks approximately. Xu et al. [21] utilized the Minhash method to filter out dissimilar routes in advance and reduce computational costs and communication overheads. Xie et al. [22] utilized property-preserving hash with road network embedding to support privacy-preserving ride-matching services. Yu et al. [49] utilized homomorphic encryption to compute secure trajectory similarity to plan an agreed path and measure trajectory deviation, so as to realize safety monitoring in online ride-hailing services. However, these algorithms or systems do not consider the sharer and taxi driver may also be the adversary to jeopardize the user's sensitive information during the whole process of ORHN, and the IRDP Ride system is the first system that considers protecting the initial point, the destination as well as the route simultaneously. As there are still a lot of algorithms or systems similar to our work, it is difficult to elaborate on all of them in just a limited space in this paper. In this section, we just discuss several typical algorithms or systems of each type, and their main ideas of them are also briefly mentioned. In order to further elaborate on these algorithms and systems, in this section, we utilize Table 2 to show the classification, the strategy, the concerned problem, and other features of each type of technology, and then we discuss the advantages and the disadvantages of them.

TABLE II
THE COMPARISON OF VARIOUS TECHNOLOGIES

The type technology	The classification	The representative systems	The strategy	The focused problem	The advantage	The disadvantage
Systems in the category of privacy preservation for requirement	Privacy preservation for snapshot query	Reference [26, 27]	Attribute generalization	Attribute differential identify	Less probability of attribute identify	Difficult to generalize all attributes
		Reference [31]	Cryptography	Information leakage	No information leakage	Long time used in encryption and decryption
		Reference [4]	Blockchain	Un-trusted of center server	Collaborate distribution calculation	High consumption and risk in collaboration
	Privacy	Reference [32,	Diversity and	The identify of	Lower probability of	The limitation of

	preservation for continuous query	33]	generalization	unity PoIs	correlation	available attributes
		Reference [34]	Service caches	Less information interaction	Less information leakage	Less trusted of cache provider
		Reference [35]	Cryptography	No information leakage	No information leakage	Long time used in encryption and decryption
		Reference [36]	Differential privacy	Generalization indistinguishable	Strong privacy assumption	Needs a large number of noises
	Privacy preservation for road network and navigation	Reference [38]	Pseudonym exchange	No tracking	Difficult to correlate pseudonyms	Other attributes can be used to track the user
		Reference [39]	Cryptography	Information leakage	No information leakage	Long time used in encryption and decryption
		Reference [41]	Differential privacy	Generalization indistinguishable	Strong privacy assumption	Needs a large number of noises
	Privacy preservation for other applications	Reference [42, 45]	Routing shifting	Path indistinguishable	Path protection	Less similar path can be used
		Reference [25, 43]	Location generalization	Location unidentified	Special application	Mainly used for special application
Systems in the category of privacy preservation for service		Reference [44]	Blockchain	Un-trusted of center server	Collaborate distribution calculation	High consumption and risk in collaboration
		Reference [20]	Cryptography	Shortest distances between riders and drivers	Securely and efficiently	Less secure for ride-matching and routes
		Reference [21]	Minhash method	Filter out dissimilar routes	Reduce computational costs and communication overheads	Without consider the taxi driver and sharer
	Privacy presersvation for online ride-hailing service	Reference [22]	Property-preserving hash with road network embedding	Privacy-preserving ride-matching	Secure ride-matching	Without consider the taxi driver and sharer
		Reference [49]	Homomorphic encryption	Compute secure trajectory similarity	Safety monitoring in online ride-hailing services	Long time used in encryption and decryption
		the IRDP Ride system	Concentric circles and multi-destination navigation	Protecting the initial point, the destination as well as the route	Consider the sharer and taxi driver may also be the adversary	Additional time in moving and road shifting

VII. Conclusion and Future Work

In order to cope with the privacy problem in ORHN, in this paper, we propose a system called the IRDP Ride system. In this system, four algorithms used to resist two potential attacks are elaborated (such as the range attack as well as the segment direction attack). The basic idea and conception of these algorithms are concentric circles and multi-destination navigation, and they are used in finding the equidistant uncertainty position and publishing the result of shifted navigation. Then the security and stability of these conceptions are analyzed with theoretical analysis and experiments, and the results will further demonstrate the superiority of the IRDP Ride system.

Although the IRDP Ride system can provide a better privacy preservation service for the user in ORHN, this system is mainly designed by the strategy of location shifting and range generalization. As a result, the time efficiency for ORHN service is a bit lower, because the user has to move to the specified pick-up point and gets off at the specified drop-off point. At the same time, the shifted navigation also occupies additional time for the taxi's moving. Furthermore, as the privacy preservation of this system is determined by a number of PoIs to generalize the initial point and destination, this system is better to be used in the area of cities with more PoIs, but difficult to be used in the countryside with fewer PoIs. Therefore, future work will be focused on how to solve the restriction of time efficiency and the restriction of PoI distribution.

References

[1] R. Wazirali, "A Review on Privacy Preservation of Location-Based Services in Internet of Things," *Intelligent Automation and Soft Computing,* vol. 31, no. 2, pp. 767-779, 2022.

[2] M. Gruteser, and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking." pp. 31-42.

[3] B. Niu, Y. H. Chen, Z. B. Wang, F. H. li, B. Y. Wang, and H. Li, "Eclipse: Preserving Differential Location Privacy Against Long-Term Observation Attacks," *Ieee Transactions on Mobile Computing,* vol. 21, no. 1, pp. 125-138, Jan, 2022.

[4] L. Zhang, D. Liu, M. Chen, H. Li, C. Wang, Y. Zhang, and Y. Du, "A user collaboration privacy protection scheme with threshold scheme and smart contract," *Information Sciences,* vol. 560, pp. 183-201, 2021/06/01/, 2021.

[5] I. Ullah, M. A. Shah, A. Khan, C. Maple, A. Waheed, and G. Jeon, "A Distributed Mix-Context-Based Method for Location Privacy in Road Networks," *Sustainability,* vol. 13, no. 22, pp. 12513, Nov, 2021.

[6] S. Shaham, M. Ding, B. Liu, S. P. Dang, Z. H. Lin, and J. Li, "Privacy Preservation in Location-Based Services: A Novel Metric and Attack Model," *Ieee Transactions on Mobile Computing,* vol. 20, no. 10, pp. 3006-3019, Oct, 2021.

[7] F. Wang, H. Zhu, X. Liu, R. Lu, F. Li, H. Li, and S. Zhang, "Efficient and Privacy-Preserving Dynamic Spatial Query Scheme for Ride-Hailing Services," *IEEE Transactions on Vehicular Technology,* vol. 67, no. 11, pp. 11084-11097, 2018.

[8] H. Shen, M. W. Zhang, H. Wang, F. C. Guo, and W. Susilo, "A Lightweight Privacy-Preserving Fair Meeting Location Determination Scheme," *Ieee Internet of Things Journal,* vol. 7, no. 4, pp. 3083-3093, Apr, 2020.

[9] Y. C. Wu, K. Wang, R. Y. Guo, Z. L. Zhang, D. Zhao, H. Chen, and C. P. Li, "Enhanced Privacy Preserving Group Nearest Neighbor Search," *Ieee Transactions on Knowledge and Data Engineering,* vol. 33, no. 2, pp. 459-473, Feb, 2021.

[10] M. Li, L. H. Zhu, and X. D. Lin, "Privacy-Preserving Traffic Monitoring with False Report Filtering via Fog-Assisted Vehicular Crowdsensing," *Ieee Transactions on Services Computing,* vol. 14, no. 6, pp. 1902-1913,

Nov, 2021.
[11] X. Y. Wang, J. F. Ma, Y. B. Miao, X. M. Liu, D. Zhu, and R. H. Deng, "Fast and Secure Location-Based Services in Smart Cities on Outsourced Data," *Ieee Internet of Things Journal,* vol. 8, no. 24, pp. 17639-17654, Dec, 2021.
[12] Z. B. Xiong, Z. P. Cai, Q. L. Han, A. Alrawais, and W. Li, "ADGAN: Protect Your Location Privacy in Camera Data of Auto-Driving Vehicles," *Ieee Transactions on Industrial Informatics,* vol. 17, no. 9, pp. 6200-6210, Sep, 2021.
[13] A. Bhardwaj, A. A. Mohamed, M. Kumar, M. Alshehri, and A. Abugabah, "Real-time Privacy Preserving Framework for Covid-19 Contact Tracing," *Cmc-Computers Materials & Continua,* vol. 70, no. 1, pp. 1017-1032, 2022.
[14] Y. Y. He, J. B. Ni, B. Niu, F. H. Li, and X. M. Shen, "Privbus: A privacy-enhanced crowdsourced bus service via fog computing," *Journal of Parallel and Distributed Computing,* vol. 135, pp. 156-168, Jan, 2020.
[15] A. Shusterman, C. Finkelstein, O. Gruner, Y. Shani, and Y. Oren, "Cache-based characterization: A low-infrastructure, distributed alternative to network-based traffic and application characterization," *Computer Networks,* vol. 200, pp. 108550, Dec, 2021.
[16] L. Zhang, M. Chen, D. Liu, L. He, C. Wang, Y. Sun, and B. Wang, "A ε-sensitive indistinguishable scheme for privacy preserving," *Wireless networks,* vol. 26, no. 07, pp. 5013–5033, 2020.
[17] L. Zhang, M. N. Chen, D. S. Liu, and J. Li, "Moving without association: an association shifting scheme for protecting destination," *Journal of Ambient Intelligence and Humanized Computing,* no. 2021, 2021.
[18] Y. Khazbak, J. Y. Fan, S. C. Zhu, and G. H. Cao, "Preserving Personalized Location Privacy in Ride-Hailing Service," *Tsinghua Science and Technology,* vol. 25, no. 6, pp. 743-757, Dec, 2020.
[19] H. N. Yu, H. L. Zhang, X. Z. Yu, X. J. Du, and M. Guizani, "PGRide: Privacy-Preserving Group Ridesharing Matching in Online Ride Hailing Services," *Ieee Internet of Things Journal,* vol. 8, no. 7, pp. 5722-5735, Apr, 2021.
[20] Y. Luo, X. Jia, S. Fu, and M. Xu, "pRide: Privacy-Preserving Ride Matching Over Road Networks for Online Ride-Hailing Service," *IEEE Transactions on Information Forensics and Security,* vol. 14, no. 7, pp. 1791-1802, 2019.
[21] Q. Xu, H. Zhu, Y. Zheng, J. Zhao, R. Lu, and H. Li, "An Efficient and Privacy-Preserving Route Matching Scheme for Carpooling Services," *IEEE Internet of Things Journal,* vol. 9, no. 20, pp. 19890-19902, 2022.
[22] H. Xie, Y. Guo, and X. Jia, "A Privacy-Preserving Online Ride-Hailing System Without Involving a Third Trusted Server," *IEEE Transactions on Information Forensics and Security,* vol. 16, pp. 3068-3081, 2021.
[23] F. H. Li, X. Y. Wang, B. Niu, H. Li, C. Li, and L. H. Chen, "Exploiting location-related behaviors without the GPS data on smartphones," *Information Sciences,* vol. 527, pp. 444-459, Jul, 2020.
[24] J. X. Huang, Y. C. Luo, S. J. Fu, M. Xu, and B. W. Hu, "pRide: Privacy-Preserving Online Ride Hailing Matching System With Prediction," *Ieee Transactions on Vehicular Technology,* vol. 70, no. 8, pp. 7413-7425, Aug, 2021.
[25] J. Kang, D. Steiert, D. Lin, and Y. J. Fu, "MoveWithMe: Location Privacy Preservation for Smartphone Users," *Ieee Transactions on Information Forensics and Security,* vol. 15, pp. 711-724, 2020.
[26] G. Sun, D. Liao, H. Li, H. F. Yu, and V. Chang, "L2P2: A location-label based approach for privacy preserving in LBS," *Future Generation Computer Systems,* vol. 74, no. 2017, pp. 375-384, Sep, 2017.
[27] J. Hua, W. Tong, F. Xu, and S. Zhong, "A Geo-Indistinguishable Location Perturbation Mechanism for Location-Based Services Supporting Frequent Queries," *Ieee Transactions on Information Forensics and Security,* vol. 13, no. 5, pp. 1155-1168, May, 2018.
[28] C. Y. Yin, X. K. Ju, Z. C. Yin, and J. Wang, "Location recommendation privacy protection method based on location sensitivity division," *Eurasip Journal on Wireless Communications and Networking,* vol. 2019, no. 1, Dec, 2019.
[29] W. Li, B. Niu, J. Cao, Y. Luo, and H. Li, "A personalized range-sensitive privacy-preserving scheme in LBSs," *Concurrency and Computation: Practice and Experience,* vol. 32, no. 5, pp. e5462, 2020.
[30] Y. Guo, H. Xie, C. Wang, and X. Jia, "Enabling Privacy-Preserving Geographic Range Query in Fog-Enhanced IoT Services," *IEEE Transactions on Dependable and Secure Computing,* vol. 19, no. 5, pp. 3401-3416, 2022.
[31] Y. L. Han, S. S. Zhu, Y. Li, and X. Lin, "APPLSS: Adaptive Privacy Preserved Location Sharing Scheme Based on Attribute-Based Encryption," *China Communications,* vol. 18, no. 3, pp. 105-121, Mar, 2021.
[32] A. Y. Ye, Y. Li, and L. Xu, "A novel location privacy-preserving scheme based on l-queries for continuous LBS," *Computer Communications,* vol. 2017, no. 98, pp. 1-10, Jan, 2017.
[33] M. Sepahkar, and M. Khayyambashi, "A novel collaborative approach for location prediction in mobile networks," *Wireless Networks,* vol. 24, no. 1, pp. 283-294, Jan, 2018.
[34] 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 Generation Computer Systems,* vol. 94, pp. 40-50, 2019/05/01/, 2019.
[35] C. Kalaiarasy, N. Sreenath, and A. Amuthan, "An effective variant ring signature-based pseudonym changing mechanism for privacy preservation in mixed zones of vehicular networks," *Journal of Ambient Intelligence and Humanized Computing,* vol. 11, no. 4, pp. 1669-1681, Apr, 2020.
[36] F. Deldar, and M. Abadi, "A differentially private location generalization approach to guarantee non-uniform privacy in moving objects databases," *Knowledge-Based Systems,* vol. 225, Aug, 2021.
[37] Y. Cao, Y. H. Xiao, L. Xiong, L. Q. Bai, and M. Yoshikawa, "Protecting Spatiotemporal Event Privacy in Continuous Location-Based Services," *Ieee Transactions on Knowledge and Data Engineering,* vol. 33, no. 8, pp. 3141-3154, Aug, 2021.
[38] S. Wang, N. Yao, N. Gong, and Z. Gao, "A trigger-based pseudonym exchange scheme for location privacy preserving in VANETs," *Peer-to-Peer Networking and Applications,* vol. 11, no. 3, pp. 548-560, 2018.
[39] S. M. Yang, S. H. Tang, and X. Zhang, "Privacy-preserving k nearest neighbor query with authentication on road networks," *Journal of Parallel and Distributed Computing,* vol. 134, pp. 25-36, Dec, 2019.
[40] X. H. Li, H. J. Zhang, Y. B. Ren, S. Q. Ma, B. Luo, J. Weng, J. F. Ma, and X. M. Huang, "PAPU: Pseudonym Swap With Provable Unlinkability Based on Differential Privacy in VANETs," *Ieee Internet of Things Journal,* vol. 7, no. 12, pp. 11789-11802, Dec, 2020.
[41] M. Khodaei, and P. Papadimitratos, "Cooperative Location Privacy in Vehicular Networks: Why Simple Mix Zones are Not Enough," *Ieee Internet of Things Journal,* vol. 8, no. 10, pp. 7985-8004, May, 2021.
[42] Z. Peng, C. Hu, C. Di, L. Hao, and L. Qi, "ShiftRoute: Achieving Location Privacy for Map Services on Smartphones," *IEEE Transactions on Vehicular Technology,* vol. 67, no. 5, pp. 4527 - 4538, 2018.
[43] I. Natgunanathan, A. Mehmood, Y. Xiang, L. X. Gao, and S. Yu, "Location Privacy Protection in Smart Health Care System," *Ieee Internet of Things Journal,* vol. 6, no. 2, pp. 3055-3069, Apr, 2019.
[44] S. H. Zou, J. W. Xi, H. G. Wang, and G. A. Xu, "CrowdBLPS: A Blockchain-Based Location-Privacy-Preserving Mobile Crowdsensing System," *Ieee Transactions on Industrial Informatics,* vol. 16, no. 6, pp. 4206-4218, Jun, 2020.
[45] Y. Wang, Z. Cai, X. Tong, G. Yang, and G. Yin, "Truthful incentive mechanism with location privacy-preserving for mobile crowdsourcing systems " *Computer Networks,* vol. 135, pp. 32-43, 2018.
[46] B. Zhao, X. Liu, W. N. Chen, W. Liang, X. Zhang, and R. H. Deng, "PRICE: Privacy and Reliability-Aware Real-Time Incentive System for Crowdsensing," *IEEE Internet of Things Journal,* vol. 8, no. 24, pp. 17584-17595, 2021.
[47] B. Zhao, S. Tang, X. Liu, and X. Zhang, "PACE: Privacy-Preserving and Quality-Aware Incentive Mechanism for Mobile Crowdsensing," *IEEE Transactions on Mobile Computing,* vol. 20, no. 5, pp. 1924-1939, 2021.
[48] V. Sadhu, S. Zonouz, V. Sritapan, and D. Pompili, " CollabLoc: Privacy-Preserving Multi-Modal Collaborative Mobile Phone Localization," *Ieee Transactions on Mobile Computing,* vol. 20, no. 1, pp. 104-116, Jan, 2021.
[49] H. Yu, H. Zhang, X. Jia, X. Chen, and X. Yu, "pSafety: Privacy-Preserving Safety Monitoring in Online Ride Hailing Services," *IEEE Transactions on Dependable and Secure Computing,* pp. 1-1, 2021.

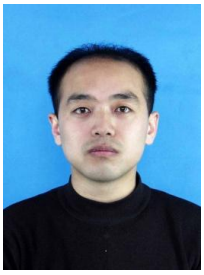
Lei Zhang was born in 1982, received the B.E., M.E. in College of Information Science and Electronic Technology at Jiamusi University, Jiamusi, China, in 2005 and 2011, respectively. He received the PH.D in 2018 in College of Computer Science and Technology at Harbin

Engineering University. He is also a professor and the director of computer software teaching and research office in College of Information Science and Electronic Technology at Jiamusi University. His research interests are security and privacy in vehicle networks, and mobile privacy protocol.

Shiyi Lin, born in 1999, received the B.E. in College of Electronic Information Engineering at Jiaying University, Meizhou, China, in 2020. She is currently a postgraduate Student in College of Information Science and Electronic Technology at Jiamusi University. Her main research interests include security and privacy in blockchain networks, smart contract, and cryptographic protocol

Chao Wang was born in 1979, received a bachelor's degree from the School of Information Science and Electronic Technology of Jiamusi University in 2002. In 2011, he received his master's degree in computer science and technology, Heilongjiang University, Harbin, China. He is also a lecturer and deputy director of the Computer Public Teaching Department and Research Office of the School of Information Science and Electronic Technology of Jiamusi University. His research interests are security and privacy in vehicular networks and mobile privacy protocols.

Jing Li was born in 1968, she is an professor in College of Information and Electronic Technology at Jiamusi University, Jiamusi, China, Her research interests are machine learning and data privacy.



Yi Liu was born in 1979, is a lecturer and master in College of Information and Electronic Technology at Jiamusi University, Jiamusi, China. His researches include image data processing and Iris image recognition.

Yue Sun was born in 1995, received the B.E., M.E. in College of Information Science and Electronic Technology at Jiamusi University, Jiamusi, China, in 2017 and 2020, respectively. She is also a teacher and a lecturer of department of robotics engineering teaching and research office in College of Information Science and Electronic Technology at Jiamusi University. She is a member of the CCF. Her research interests are security and privacy in mobile social networks.