IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 8, 2019, accepted June 24, 2019, date of publication July 5, 2019, date of current version July 24, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2927140

RcDT: Privacy Preservation Based on R-Constrained Dummy Trajectory in Mobile Social Networks

JINQUAN ZHANG[1], XIAO WANG[1], YANFENG YUAN[1], AND LINA NI[1,2]
[1]College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China
[2]Key Laboratory of the Ministry of Education for Embedded System and Service Computing, Tongji University, Shanghai 201804, China

Corresponding author: Lina Ni (nln2004@163.com)

This work was supported in part by the National Key Research and Development Programs Project of China under Grant 2017YFC0804406, in part by the NSF of China under Grant 61672321, Grant 61771289, Grant 61832012, and Grant 61373027, in part by the Training Program of the Major Research Plan of NSF of China under Grant 91746104, in part by the Project of Shandong Province Higher Educational Science and Technology Program under Grant J15LN19, and in part by the Open Project of Tongji University Embedded System and Service Computing of Ministry of Education of China under Grant ESSCKF 2015-02.

ABSTRACT The boom of mobile devices and location-based services (LBSs) greatly enriches the mobile social network (MSN) applications, which bring convenience to our daily life and, meanwhile, raise serious privacy concerns due to the potential disclosure risk of location privacy. Besides the single-location privacy, trajectory privacy is another important type for location privacy leakage. In this paper, focusing on the trajectory privacy preservation in MSNs, we propose a privacy preservation scheme based on the radius-constrained dummy trajectory (RcDT) in MSNs. Particularly, by constraining the generated circular range with radius R for the location where a user sends LBS requests, we present the radius-constrained dummy location (RcDL) algorithm to generate the dummy location set of the user's real location. Furthermore, based on the generated dummy locations, we put forward the RcDT algorithm to generate the dummy trajectory set that has higher similarity to the real trajectory comprehensively considering the constraints of both the single-location exposure risk and trajectory exposure risk. Thus, the user's trajectory privacy preservation in MSNs is enhanced since the possibility of identifying users' real trajectories and malicious attacks are reduced. The simulation results demonstrate that our RcDT scheme can have better performance and privacy degree than the existing methods.

INDEX TERMS Privacy preservation, mobile social networks, trajectory privacy, location-based service.

I. INTRODUCTION

The proliferation of mobile devices furnished with abundant location-based services (LBSs) has made the location-aware applications increasingly extensive [1]–[3]. Applying LBSs to social networks has spawned mobile social networks (MSNs) that allow users to discover potential friends around them, share information with each other and check-in to a mall site to obtain special discount information, etc. [4]–[7]. However, these services may trigger the serious privacy concern due to the potential disclosure risk of location privacy. Therefore, the research of location privacy preservation technique has aroused much concern [8]–[10].

In real life, mobile users of LBSs do not merely stagnate in a single location point, while their locations and time may constantly change. The location trajectory is a temporal-spatial sequence formed by multiple consecutive queries issued by different users at different times in mobile state, that is, the location trajectory is formed by a sequential time series of single location points [11]. Hence, trajectory privacy preservation is also the salient issue that location privacy protection should address [12]. These trajectories which comprise abundant users' background knowledge have more attraction for attackers to mine valuable information, such as users' life style, relationships between people and other personal private information, etc. Therefore, it is a more stern challenge to protect the trajectory privacy in MSNs [13]–[17].

Incorporating the location privacy preservation techniques into the characteristics of trajectory privacy, some approaches of trajectory privacy preservation have been

The associate editor coordinating the review of this manuscript and approving it for publication was Sajjad A. Madani.

proposed ([18]–[20]), such as dummy trajectory, trajectory suppression, trajectory k-anonymity, etc. Mixing fake trajectories into real trajectories, the dummy trajectory technique [21] is employed to obscure the real trajectories, which is hard for attackers to distinguish users' real trajectories. Thus, users' preferences and behavior privacy cannot be deduced so as to realize trajectory privacy protection. Trajectory suppression technique [22] hides some of the sensitive positions to realize trajectory privacy protection. However, excessive suppression may cause serious loss of trajectory information and degrade service quality. Trajectory k-anonymity technique [23] forms the information of k trajectories, which is derived from the location k-anonymity. However, it is tough to determine the value of k and select the $k-1$ trajectories for trajectory k-anonymity.

In order to further solve the problems of single location protection and trajectory privacy preservation in MSNs, in this paper, we present a privacy preserving scheme based on Radius-constrained (R-constrained for short thereafter) dummy trajectory (RcDT) in MSNs. Around each user's real location point, the user randomly sends queries to generate a location within a specific circular range with radius R, namely a dummy location. If the generated dummy location meets certain requirements, it is placed into the dummy location set. Otherwise, it is regenerated until the number of generated dummy locations meet users' requirements. Whereafter, incorporated single location exposure risk (SE) and trajectory exposure risk (TE), the dummy trajectory is generated based on these dummy locations. It is noted that single location privacy protection may generally focus on timeliness, while the trajectory privacy protection does not pay attention to real-time performance which instead highlights users' background knowledge. Our scheme considers both the timeliness and users' background knowledge jointly and can effectively draw on each other's strengths.

Compared with existing methods, our scheme restrains the generation range of the locations with a radius R around the real location to constrain the dummy locations and trajectories, this is why we call our scheme R-constrained dummy trajectory, which makes the generated dummy trajectory has a higher similarity with the real one. Therefore, it is difficult for attackers to identify real trajectories. Single location exposure risk and trajectory exposure risk are evaluated respectively, meanwhile the dummy location that easily exposes user's privacy is deleted, ensuring the protection for the trajectory privacy of mobile users.

The main contributions of this paper are summarized below.

1) We propose a R-constraint dummy trajectory privacy preservation scheme (RcDT) in MSNs. We first present the R-constrained dummy location (RcDL) algorithm to generate the set of dummy locations of a user's real location. Then we put forward RcDT algorithm to generate a set of dummy trajectories employing the generated dummy locations comprehensively considering SE and TE.
2) We analyze the superiority and privacy degree of RcDT. Through the constraint of SE and TE, the similarity between the generated dummy trajectories and the real ones is higher, reducing the possibility of identifying real trajectories for attackers. Therefore, the level of trajectory privacy preservation is increased.
3) We conduct extensive simulations to verify RcDT. Simulation results demonstrate that our RcDT scheme can provide more enhanced trajectory privacy preservation than the existing methods.

The rest of the paper is organized as follows. Section II reviews the related work. In Section III, we give the preliminary knowledge related to RcDT. The system model and design motivation of RcDT are given in Section IV. Following that, in Section V, the R-constrained dummy trajectory scheme is proposed. Simulation experiments are given in Section VI. Finally, we draw our conclusions and give the future work in Section VII.

II. RELATED WORK

Trajectory privacy preservation has been attracting the attention from both academia and industry. This issue draws even more attention in the recent years due to the booming of LBSs.

k-anonymity technique also has a good application in trajectory privacy preservation. Huang *et al.* [23] proposed a k-anonymity method utilizing users' personalized privacy parameters setting such that anonymous servers generalized users' identification information and location information as much as possible in trajectory protection. An r-anonymous mechanism was proposed in [24], which sends a query request by randomly selecting a location from a set of locations formed by all r trajectories, so that it is difficult for attackers to identify users' real trajectories.

In [22], Terrovitis and Mamoulis used suppression method for trajectory privacy protection, which does not release the trajectory data with high sensitive degree or visiting frequency. With respect to choose the appropriate suppression points as breakthrough points, Zhao et al. put forward a trajectory frequency suppression method [25] to interfere trajectory data with fake data and suppress the released trajectory data as a whole.

Aiming at the privacy threaten aroused by the check-in function in social applications based on LBSs, Huo et al. proposed a private CheckIn trajectory privacy protection method [26] which realizes trajectory data anonymity by constructing prefix tree. Yang et al. put forward a personalized check-in service data dissemination algorithm [27] based on locations to confuse the sign-in data of users. Furthermore, in [28], Hossain *et al.* removed the validity of location information from trajectory data to protect location privacy and infer the trajectory privacy attack.

You et al. proposed one trajectory preservation scheme [29] using stochastic and rotation methods to generate dummy trajectories similar to user's real ones making the trajectory difficult to be distinguished. Huo *et al.* [30] generated dummy

trajectories by producing dummy data to increase sampling points on the basis of the real trajectory sampling points, reducing the recognition rate of users' real trajectories. Although this method is relatively simple with a small amount of calculation, when the degree of privacy protection required is relatively high, the volume of interference data introduced will be very large so that the procreant dummy data will occupy plenty of storage space.

In [31], we have proposed a preliminary trajectory privacy preserving method by constraining the range of the dummy trajectories in MSNs. However, this method neglects the design motivation, in-depth algorithm design and analysis. What's more, this method does not carry out simulations to verify its performance.

Existing research on trajectory privacy preservation of LBSs mostly focuses on how to store and index offline trajectory data. In this paper, we propose a privacy preserving method based on R-constrained dummy trajectory (RcDT) in MSNs. We restrain the generation range of the locations, which makes the generated dummy trajectories have a higher similarity with the real one, so that it is difficult for the attackers to identify real trajectories.

III. PRELIMINARIES

In this section, we give the preliminary knowledge of our RcDT scheme needed.

A. DUMMY TRAJECTORY GENERATION TECHNIQUES

Dummy trajectory generation technique, whose basic idea is to confuse users' real trajectories by pretending the dummy ones, is a trajectory privacy protection method corresponding to the dummy location technique [20], [32]. The higher the similarity between the generated dummy trajectories and real ones, the better the privacy protection effects. Therefore, how to make generated dummy trajectories similar to real ones is the key concern of this method. Existing dummy trajectory generation techniques mainly include stochastic generation method [33] and rotation generation method [34].

1) STOCHASTIC GENERATION METHOD [33]

According to the real trajectories within a period of time, the start location point and end one are found out, and the dummy location points are stochastically generated respectively in the moving area of user's real trajectory. Then the dummy starting point and end point are connected whose run time is ensured to be same as that of the real trajectories. This method is simple to implement but does not consider the distance deviation offset. If the generated dummy trajectory deviates greatly from the real one, it will not achieve the purpose of confusion, thus the attacker can easily identify the real or dummy trajectories, resulting in user information exposure.

2) ROTATION GENERATION METHOD [34]

Users' real trajectories are selected as benchmark, a location point is randomly selected on the trajectory as the rotation point. Then, the dummy trajectory is generated by rotating the real trajectory at a certain angle. If the rotation angle is suitable, the dummy trajectory generated may close to the real one. Otherwise, the deviation offset of the dummy trajectory may be relatively larger. Besides, only after users' trajectories are formed can the dummy ones be generated in this method. Thus, the attackers can easily distinguish the real and dummy trajectories after analyzing their generation time due to the larger difference of time similarity between the generated dummy trajectories and the real ones.

B. RELEVANT DEFINITIONS

In our scheme, we determine the effect of the generated dummy trajectories based on the distance between the dummy and real trajectories, location points and attackers' identification probability of trajectories during the process of generating the dummy trajectories. In order to facilitate the algorithm description, we present the following definitions.

Definition 1 (Location distance (L_{dist})): Let $RL = (x, y)$ be a real trajectory location point, $DL = (x', y')$ be a dummy trajectory location point of RL. The *location distance* between RL and DL is defined as the Euclidean distance between RL and DL, denoted by $L_{dist}(RL, DL)$, that is,

$$L_{dist}(RL, DL) = \sqrt{(x - x')^2 + (y - y')^2}. \tag{1}$$

Definition 2 (Trajectory distance (T_{dist})): Trajectory is formed the collection of n points of location coordinates within a period of time. Let RT be a real trajectory whose dummy trajectory is DT. The *trajectory distance* between RT and DT is defined as the average distance of all the location points corresponding to RT and DT at the same time instant, denoted by $T_{dist}(RT, DT)$, that is,

$$T_{dist}(RT, DT) = \frac{1}{n} \sum_{t=1}^{n} L_{dist}(RL^t, DL^t), \tag{2}$$

where RL^t and DL^t are users' real location and dummy location at the time instant t respectively.

Definition 3 (Single location exposure risk (SE)): Suppose that a user's real trajectory RT contains n location points, D_i is the set of location points containing the real and dummy locations of the user at a certain location point. Then the probability of exposing a single location from a set of location points is $\frac{1}{|D_i|}$. The user's *single location exposure risk* is defined as the probability of exposing the real location from all the sets of location points, denoted by SE, that is,

$$SE = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{|D_i|}. \tag{3}$$

Definition 4 (Trajectory exposure risk (TE)): Suppose that the number of dummy trajectories formed by a user's real trajectory is m, where there is at least one intersection between k trajectories and no intersection among the other $m-k$ trajectories. Let the number of the trajectories from the starting point to the end point formed by these k overlapping trajectories be T_k. Then the *trajectory exposure risk* of the user

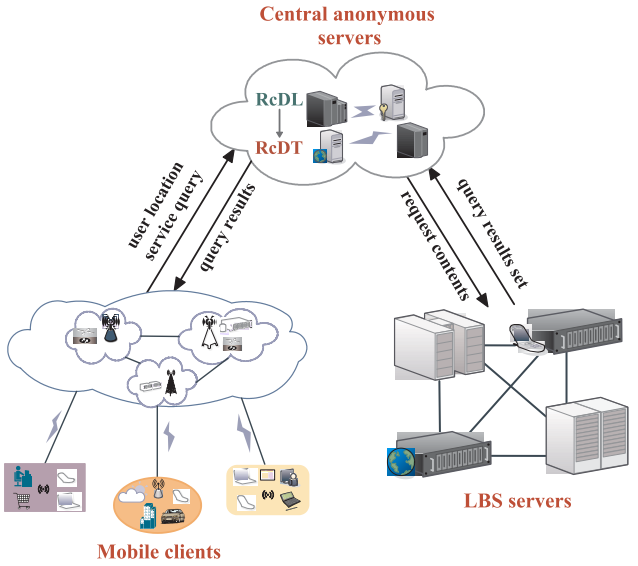


FIGURE 1. Architecture of system model.

is defined as the probability of exposing the real trajectories from all the possible trajectories, denoted by TE, that is,

$$TE = \frac{1}{(m-k)+T_k}. \tag{4}$$

Definition 5 (Distance deviation (DD)): The *distance deviation* is defined as the mean value of the offset distance of the location points from all the dummy trajectories to the real trajectories at the time instant t, denoted by DD, that is,

$$DD = \frac{1}{n}\sum_{t=1}^{n}(\frac{1}{m}\sum_{j=1}^{m}L_{dist}(RL^t, DL_j^t), \tag{5}$$

where $L_{dist}(RL^t, DL_j^t)$ is the distance between each location point of real and dummy trajectories at each time instant t, m is the number of dummy trajectories.

Definition 6 (Distance deviation degree (DD_{degree}): The *distance deviation degree* is defined as the ratio of the mean value of DD versus the radius $|R|$ of the circular range generated by dummy locations, denoted by DD_{degree}, that is,

$$DD_{degree} = \frac{\frac{1}{n}\sum_{i=1}^{n}DD_i}{|R|} \times 100\%. \tag{6}$$

IV. SYSTEM MODEL AND DESIGN MOTIVATION

In this section, we first put forward the architecture of our system model, and then give the design motivation of RcDT.

A. SYSTEM MODEL

We consider a system model that has three components: mobile clients, central anonymous servers and LBS location servers as shown in Fig. 1. Central anonymous servers are the trusted third party of location privacy protection.

The communication between the mobile clients and the central anonymous servers is cryptographically authenticated, assuming that it is hard to be intruded by an attacker. While the communication between the central anonymous servers and the LBS servers is untrusted and here it is assumed that attackers can eavesdrop or intercept the communication information. The main workflow of our system is as follows:

1) The mobile clients (users) issue LBS requests at a certain location, meanwhile their location information and privacy protection parameters are sent to the central anonymous servers.
2) The central anonymous servers generate dummy locations according to the privacy parameters sent by the clients employing RcDL algorithm, then the request contents in the real and dummy locations are sent to the LBS servers together.
3) After the LBS servers receive the requests, all the results related to the real and the dummy locations are queried and then sent to the central anonymous servers.
4) The central anonymous servers obtain the relatively accurate query results based on the real locations of the clients, sending the refinement results back to the clients.
5) After the above steps, a single location service request has been completed. The privacy preserving trajectory formed by multiple location service requests using RcDT algorithm is the output of our system.

In fact, trajectory privacy protection is also a special location privacy protection, thus single location privacy protection needs to be taken into account in trajectory privacy protection. Single location privacy protection is generally real time, while the trajectory privacy protection does not pay attention to real time effects which instead highlights attackers' background knowledge. Our system considers both the privacy protection ideas jointly and can effectively draw on each other's strengths.

B. DESIGN MOTIVATION

Dummy location technique, often used in location privacy protection, is a simple and effective privacy preserving technique. The so-called dummy location technique is to utilize dummy location (x', y') instead of users' real location (x, y) to send requests and obtain corresponding location services.

Similarly, the dummy trajectory technique can be used in trajectory privacy protection. As mentioned before, the main idea of dummy trajectory technique is that based on users' location trajectories, fake information is added into real trajectories to generate some dummy trajectories according to certain strategies, reducing the identification risk of users' real trajectories. The closer to the real trajectories, the more confusion of the dummy trajectories.

For example, as can be seen in Table 1, the original location information of users U_1, U_2 and U_3 at the time instant t_1, t_2 and t_3 is recorded, and three user trajectories are formed respectively.

TABLE 1. Example of Original Data of Users' Trajectories.

Time / User	t_1	t_2	t_3
U_1	$(1,1)$	$(4,2)$	$(5,5)$
U_2	$(1,3)$	$(3,2)$	$(6,2)$
U_3	$(2,1)$	$(3,4)$	$(6,6)$

Then the mixed data sets are formed by interfering the real trajectory data in Table 1 with some dummy trajectory data. The resulted dummy trajectory data are shown in Table 2. As can be seen from Table 2 that the records of U_1', U_2' and U_3' are the corresponding dummy trajectories of those of U_1, U_2 and U_3, respectively. In this way, the identification probability of each real trajectory will be $1/(k+1)$, where k is the number of the added dummy trajectories. The larger the k value, the less likely the real trajectories will be identified, and the higher the security for the user privacy.

TABLE 2. Example of Resulted Data of Users' Trajectories.

Time / User	t_1	t_2	t_3
U_1	$(1,1)$	$(4,2)$	$(5,5)$
U_2	$(1,3)$	$(3,2)$	$(6,2)$
U_3	$(2,1)$	$(3,4)$	$(6,6)$
U_1'	$(1,2)$	$(4,3)$	$(5,6)$
U_2'	$(1,4)$	$(4,2)$	$(7,3)$
U_3'	$(2,2)$	$(4,5)$	$(6,7)$

When the generated dummy trajectories are close to the real ones of users, they can have interference effects. However, if the added dummy trajectories are irrelevant to the real ones, attackers can easily find the real trajectories through analysis. With respect to how to generate dummy trajectories similar to the true trajectories to achieve users' trajectory privacy protection [35], the following factors should be considered when dummy trajectories are generated.

1) THE NUMBER OF DUMMY TRAJECTORIES

Generally, the more the number of dummy trajectories, the more difficult to identify the real ones. However, the number of the generated dummy trajectory should conform to the actual situation. For example, it is common that the amount of urban mobile users are more than that of suburb mobile users and the population of the mobile users at daytime are more than that at night. Due to different time and location, the number of the generated dummy locations may be differentiated.

Here, we put forward a function $k_t = f(Uid, RL^t)$ to generate the number of dummy locations, where Uid is a user's identification information, RL^t represents the user's real location at the time instant t and k_t is the number of dummy location generated by the real one at the time instant t.

2) THE TIME SIMILARITY BETWEEN DUMMY AND REAL TRAJECTORIES

Although the run time of the generated dummy trajectories doesn't have to be as real-time as the location privacy protection, it should be consistent with the actual situation of real trajectories, and the variance between the dummy location generation time and users' query time can't differ too much as well. Therefore, the time similarity between the dummy and real trajectories can be measured by the following formula:

$$sim_{t,t'} = \frac{||t' - t||}{\theta}, \tag{7}$$

where t represents users' query time in real location, t' represents users' dummy location generation time, θ represents the time threshold which is set by users, $||\cdot||$ represents the norm of "$\cdot$".

3) THE SPATIAL SIMILARITY BETWEEN DUMMY AND REAL TRAJECTORIES

The resulted dummy locations should close to the real ones in space so that the generated dummy trajectories are closer to the real ones to produce interference. If the deviation between the generated dummy trajectories and real ones is larger, it is easier for attackers to identify. The spatial similarity between the dummy and real trajectories can be measured by the following formula:

$$sim_{l,l'} = \frac{|| < x', y' > - < x, y > ||}{\delta}, \tag{8}$$

where $<x, y>$ represents users' real locations, $<x', y'>$ represents users' dummy locations, δ represents the anonymous area of dummy locations.

V. R-CONSTRAINED DUMMY TRAJECTORY SCHEME

Our RcDT scheme is inspired by the stochastic generation method of dummy trajectories. The dummy locations are generated at each location point of users' queries which have certain constraints to make their deviation from the real location points not too much. In this way, the similarity between the generated dummy trajectories and the real ones is higher, making it difficult for attackers to identify and reducing the exposure risk of user trajectories at the same time.

A. THE OVERALL PROCEDURE OF RCDT

Users' real trajectories are the connection of their locations sending LBS requests at different time. Taking the example of one user sending requests, the overall procedure of users' dummy trajectories generation scheme is described as follows:

1) Obtain the coordinates of the location point A where the user requests a LBS.
2) The user's location of the requested LBS is as the center of the circle, and the circular area is formed with radius R set by the user.
3) The dummy location is randomly generated around the coordinates of the user's location point.

4) If the dummy locations are within the circular area with radius R, the dummy locations are in accordance with the requirements; otherwise, the dummy locations will be regenerated until the number of location points satisfies the user's requirements.
5) The location trajectory is formed by the dummy location points. If the dummy location points are within the range of the single location exposure risk SE and the formed dummy trajectory is within the range of the trajectory exposure risk TE, these location points are added to the location points set.
6) The dummy trajectory which is formed by the set of the location points is the one that meets the requirements.

The trajectories are generated according to the above steps to obtain a certain number of dummy trajectories which are similar to the real ones.

B. ALGORITHM DESIGN

Firstly, we initialize some parameters. Suppose that one user's real location coordinates at the time instant t are $RL^t = (x^t, y^t)$, the number of generated dummy locations for each real location is k_t, which is determined by the dummy locations generation number function $k_t = f(Uid, RL^t)$, then the set of the corresponding k_t dummy locations at the time instant t is $SL_d^t = \{DL_1^t, DL_2^t, \cdots, DL_{k_t}^t\}$, where $DL_i^t = (x_i^t, y_i^t)$ $(i = 1, 2, \cdots, k_t)$ is the dummy locations of RL^t at the time instant t. The user's LBS request at the time instant t is $SL_d^t = Q(id, RL^t, k_t)$, where $Q(\cdot)$ is a function to generate the set of the dummy locations according to the user's identification information Uid, real location RL^t and the number of generated dummy locations k_t.

Besides, the real location trajectory generated by the user's n locations sending LBS requests is represented by $RT = (RL^1, RL^2, \cdots, RL^n)$, and the dummy trajectory formed by real trajectory is represented by $DT = (DL_{p_1}^1, DL_{p_2}^2, \cdots, DL_{p_n}^n)$, where $DL_{p_i}^t (t = 1, 2, \cdots, n)$ is a dummy location of the real location RL^t at the time instant t, that is, $DL_{p_t}^t \in \{DL_1^t, DL_2^t, \cdots, DL_{k_t}^t\}$. The dummy trajectories set generated by real trajectory RT is denoted by $ST_d = \{(DL_{p_1}^1, DL_{p_2}^2, \cdots, DL_{p_n}^n) | p_t \in \{1, 2, \cdots, k_t\} (t = 1, 2, \cdots, n)\}$.

When the user requests the LBS, the area of the dummy locations is generated, namely the circular area with the true position as the center and R as the radius is generated. Meanwhile, the user needs to send corresponding parameters to the central anonymous servers including the location service requests Q, dummy location generation number at the time instant t in each location k_t $(t = 1, 2, \cdots, n)$, radius of the generated range of dummy location R, single location exposure risk SE and trajectory exposure risk TE. Central anonymous servers send the sets of the locations that are formed by the dummy and real locations to LBS servers. LBS servers return the corresponding query result set to the central anonymous servers which then send the sorted results to the users.

1) ALGORITHM ILLUSTRATION

The details of generating the set of R-constrained dummy locations (RcDL) according to the user's real location are elaborated in Algorithm 1.

In this algorithm, one dummy location is randomly generated. If its distance from the real location is within the R region, it is added to the dummy location set; otherwise, the dummy location will be regenerated until the desired dummy location set is come into being. Then, the dummy locations of other location points on the user's trajectories are created in chronological order in the same way.

Algorithm 1 : RcDL generation algorithm

Input: User's real location at the time instant t $RL^t = (x^t, y^t)$, k_t and R.
Output: Dummy locations set SL_d^t.

$SL_d^t = \phi$
for $i = 1$ to k_t **do**
  $DL^t = (x_i^t, y_i^t) = Random(RL^t)(x_i^t \in [x^t - R, x^t + R], y_i^t \in [y^t - R, y^t + R])$
  //randomly generate the dummy locations around RL^t
  $d = L_{dist}(RL^t, DL^t) = \sqrt{(x^t - x_i^t)^2 + (y^t - y_i^t)^2}$ //calculate Euclidean distance between dummy and real locations according to Eq. (1)
  if $d \le R$ **then**
    $SL_d^t = SL_d^t \cup \{DL^t\}$
  else
    $q = Random(8)$ //generate a random number within 8
    if $q \le 2$ **then**
      $x_i^t = x^t - \sqrt{R^2 - (y^t - y_i^t)^2}$
    else if $q \le 4$ **then**
      $x_i^t = x^t + \sqrt{R^2 - (y^t - y_i^t)^2}$
    else if $q \le 6$ **then**
      $y_i^t = y^t - \sqrt{R^2 - (x^t - x_i^t)^2}$
    else
      $y_i^t = y^t + \sqrt{R^2 - (x^t - x_i^t)^2}$
    end if
  end if
  $DL^t = (x_i^t, y_i^t)$
  $SL_d^t = SL_d^t \cup \{DL^t\}$
end for
return SL_d^t

Both the single location exposure risk SE and the dummy trajectory exposure risk TE should be considered in generating the set of R-constrained dummy trajectories (RcDT) according to the user's real trajectory. The detailed implementation procedure of RcDT generation algorithm is illustrated in Algorithm 2.

In this algorithm, starting from the initial location, Algorithm 1 is used to generate a dummy location set for each location. Then, the dummy trajectory set of the current location is generated combined the dummy trajectory of

previous location and the dummy location set of the current location, and the dummy trajectories that do not meet the requirements of the spatial similarity *SIM* between dummy and real trajectories are deleted according to formula (8). Afterwards, according to formulas (3) and (4), *SE* and *TE* are calculated. If *SE* and *TE* meet the requirements, the dummy location set of the current location as well as that from the initial location to the current location is obtained until the dummy trajectory set of whole real trajectories is generated; otherwise the dummy location and trajectory sets of the current location are regenerated until meet the requirements.

Algorithm 2 : RcDT generation algorithm

Input: Location service request set $RT = \{RL^1, RL^2, \cdots, RL^n\}$, $k_t(t = 1, 2, \cdots, n)$, R, SIM, SE, TE.

Output: Dummy trajectories set ST_d.

1: Call **Algorithm 1** to generate dummy locations of real location RL^1: $SL_d^1 = \{DL_1^1, DL_2^1, \cdots, DL_{k_1}^1\}$
2: $ST_d = \{(DL_1^1), (DL_2^1), \cdots, (DL_{k_1}^1)\}$
3: **for** $t = 2$ to n **do**
4: $ST_d' = \phi$
5: Call **Algorithm 1** to generate dummy locations of real location RL^t: $SL_d^t = \{DL_1^t, DL_2^t, \cdots, DL_{k_t}^t\}$
6: **for** each $(DL_{p_1}^1, DL_{p_2}^2, \cdots, DL_{p_{t-1}}^{t-1}) \in ST_d$, $p_j \in \{1, 2, \cdots, k_j\}$ $(j = 1, 2, \cdots, t-1)$ **do**
7: **for** each $DL_i^t \in SL_d^t$ $(i = 1, 2, \cdots, k_t)$ **do**
8: Calculate sim through $(RL^1, RL^2, \cdots, RL^t)$ and $(DL_{P_1}^1, DL_{P_2}^2, \cdots, DL_{P_{t-1}}^{t-1}, DL_i^t)$ according to Eq. (8)
9: **if** $(sim \leq SIM)$ **then**
10: $ST_d' = ST_d' \cup \{(DL_{p_1}^1, DL_{p_2}^2, \cdots, DL_{p_{t-1}}^{t-1}, DL_i^t)\}$
11: **end if**
12: **end for**
13: **end for**
14: Calculate SE_t and TE_t through ST_d' according to Eqs. (3) and (4)
15: **if** $(SE_t \leq SE) \&\& (TE_t \leq TE)$ **then**
16: $ST_d = ST_d'$
17: **else**
18: Goto *step* 5
19: **end if**
20: **end for**
21: **return** ST_d

In Algorithm 2, the dummy trajectories are formed according to dummy locations at time instant t within the constraints of *SE* and *TE*. Furthermore, the generated trajectory offset also meets the requirements, thus the anonymous dummy trajectories are formed which meet users' privacy demands.

2) TIME COMPLEXITY ANALYSIS OF ALGORITHMS

Theorem 5.1: Let RT be a real trajectory with n locations, $k_t(t = 1, 2, \cdots, n)$ be the number of dummy locations for each location. Denote $k_{\max} = \max\{k_t | t \in [1, n]\}$, then the time complexity to generate dummy location and trajectory sets which meet the requirements of *SE* and *TE* is $O(n^{k_{\max}})$.

Proof: Firstly, in Algorithm 1, when the dummy locations of $t^{th}(t = 1, 2, \cdots, n)$ location are generated, if R constraint is not satisfied, a random number is generated to modify x_i^t or y_i^t to satisfy the R constraint. Therefore, the time complexity of generating the dummy location set of the t^{th} location is $O(k_t)$.

Then, in Algorithm 2, when generating dummy trajectories, the number of the dummy trajectories from the initial location to the current t^{th} location is $k_1 \times k_2 \times \cdots \times k_t$ at most. According to formula (8), the time to delete dummy trajectories that do not meet the requirements is $O(k_1 \times k_2 \times \cdots \times k_t)$. According to formulas (3) and (4), the time to calculate *SE* and *TE* also is $O(k_1 \times k_2 \times \cdots \times k_t)$.

Similarly, if the requirements of the privacy exposure risks are not met, the dummy location and trajectory sets need to be reconstructed. In general, the number of the repetitions is limited. Thus the time to generate the dummy trajectory set that satisfies the constraints of *SE* and *TE* still is $O(k_1 \times k_2 \times \cdots \times k_t)$.

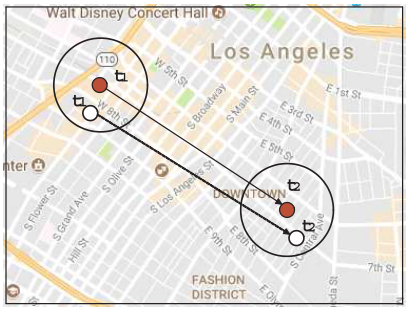
Therefore, the time complexity of generating dummy location and trajectory sets for the real trajectory RT is $O(k_1) + O(k_2) + \cdots + O(k_n) + O(k_1 \times k_2 \times \cdots \times k_n) = O(n^{k_{\max}})$. ■

C. ALGORITHM EXPLICATION

Users' location information is obtained when users query LBS at different time and then the users' real locations are connected to form trajectories. According to the traditional stochastic generation method, the dummy locations are generated randomly at every location and the obtained dummy locations and trajectories may have larger difference with the real trajectories, thus attackers could easily identify the real and dummy trajectories.

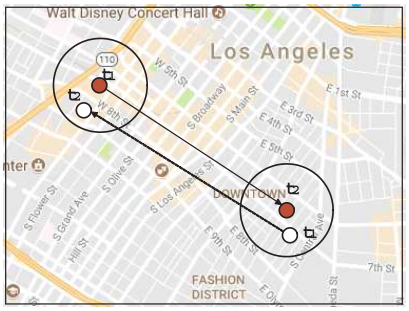
In spite of our proposed scheme also based on stochastic generation method, the generation scope of dummy locations is constrained, and the location points in the scope of *SE* are selected as the locations on the trajectories. Thus the location trajectories are constrained within a certain distance deviation (*DD*). The trajectories are formed by constraining dummy locations and selecting location points within the scope of (*SE*), making the location trajectory also constraint within a certain *DD* to improve the similarity between the dummy and real trajectories, therefore it is difficult for attackers to make identification.

As shown in Fig. 2, the solid red dots represent the target user's real locations while the hollow dots represent the dummy locations, and the solid lines represent the user's real trajectories while the dashed lines represent the user's dummy trajectories, the arrow direction is the direction of the user's trajectories. Note that the generated trajectories of Fig. 2 (a) meets the requirements of our definitions while the generated locations of Fig. 2 (b) is outside of the scope of radius R which does not conform to our requirements. What's more, In Fig. 2 (c), the direction deviation between the dummy and



(a) Meet the requirements.

(b) Unqualified.



(c) Unqualified.

FIGURE 2. Illustrations of dummy trajectories generation of RcDT.

real trajectory directions is larger, which can't conform to our requirements.

D. ALGORITHM SUPERIORITY ANALYSIS

Dummy locations generated by stochastic generation method distributes randomly, which is simple but does not consider distance deviation. If the generated dummy trajectories largely deviate from the real ones, there will be no confusion, on this occasion attackers can easily identify the real and dummy trajectories, leading to the exposure of users' information.

The dummy trajectory generation method of our RcDT scheme is based on dummy location stochastic generation idea and some improvements are made. The dummy locations constrained with user-defined radius R are generated randomly at the location points of users' queries, that is, the generated dummy locations are in a circular region of radius R, namely the distance between the dummy and real locations is no more than R. The number of dummy locations generated per location is also defined by users' sensitivity.

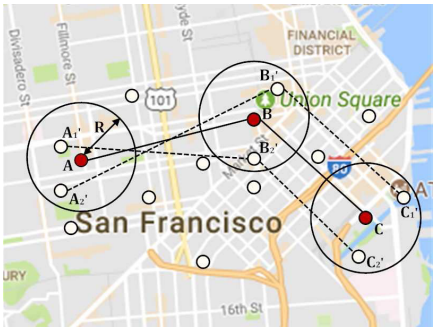


FIGURE 3. Schematic of our RcDT scheme.

The dummy locations which meet the requirements constitute the set of the dummy location points. The dummy trajectories formed by connecting the points in the set of dummy location points need to be deviated from real locations not too much while satisfying the constraints of SE and TE. The resulting dummy trajectories with interference are close to the real ones which make it difficult for attackers to identify and reduce the exposure risk of users' movement trajectories. The schematic of privacy protection method based on RcDT is shown in Fig. 3.

As shown in Fig. 3, the real trajectories of the mobile user are formed by black solid line of the real locations A, B, and C (represented by the solid red dots), where some dummy locations are generated stochastically. Only A'_1, A'_2, B'_1, B'_2, C'_1 and C'_2 (represented by the hollow dots) are within the constraint range of radius R, which could form the sets of the locations point that satisfy the constraint requirements. These point sets form several dummy trajectories. There are two dummy trajectories (represented by the dotted lines) in Fig. 3. It can be seen that it is easier to obtain the dummy trajectories that meet the requirements through the constraint of randomly generated dummy locations. Compared with the stochastic generation method, our algorithm generates a higher proportion of valid dummy trajectories.

At each location point, some dummy locations are generated in real time, which could better protect the background knowledge against the attackers. For example, when a mobile user requests a service, the format of the request information may be $Q(Uid, RL^t, k_t)$. Every time the Uids of the user's requests are the same. Note that if the user's queries are in a row, the formed trajectories will be within a very small area or even be a point. Thus, the attackers could easily obtain the real information of the user via many-time comparative analysis of the location areas of the user's requests.

Fortunately, our RcDT schema will generate dummy locations in each location point, so it is difficult for attackers to attack a single location. Even if mobile users continuous apply location services with association between locations, it is hard to be identified by attackers, thus users' privacy is effectively protected.

VI. EXPERIMENTS

In this section, we implement our proposed RcDT scheme and evaluate its performance via extensive experiments.

TABLE 3. Initial Parameters Settings.

No. \ Value	SE	TE	DD	R
1	5%	5%	5%	$5km$
2	10%	10%	10%	$5km$
3	15%	15%	15%	$5km$
4	20%	20%	20%	$5km$
5	25%	25%	25%	$5km$
6	30%	30%	30%	$5km$

A. SIMULATION ENVIRONMENT

All algorithms are implemented in Matlab 2015a and tested on a PC with Windows 10 64 bits operating system, 8GB RAM and Intel (R) Core (TM) i7-4700MQ CPU@3.4GHz processor.

In our experiment, Thomas Brinkhoff, a Network-based Generator of Moving Objects [36], is used to generate spatiotemporal data as our simulation dataset. The privacy protection parameters of the experiments are determined according to users' demand. The generated dataset consists of users' trajectory information and social networking connections.

B. SIMULATIONS AND PERFORMANCE EVALUATION

In our experiments, an area of 10 $km * 10$ km is selected as the simulation region where mobile users are deployed randomly. Furthermore, we assume that the user who initiates LBS requests is located at some point in the region and again sends the LBS requests when he reaches the next location point at each time interval *inte*.

To intensively verify the performance of our RcDT, we set six groups of values for the four parameters including the radius of the generated range *R*, the single location exposure risk (*SE*), the trajectory exposure risk (*TE*), and distance deviation (*DD*). The initial simulation parameters settings are listed in Table 3.

We compare our RcDT scheme with a stochastic location generation dummy trajectory method (RPDT) [33] and virtual rotation generation dummy trajectory method (TPPA-TD) [34] respectively based on the initial parameters settings. RPDT algorithm randomly generates dummy locations around the real locations to form dummy trajectories. However, our RcDT constrains the generation scope of the dummy locations during the formation of dummy trajectories. The trajectories generated by TPPA-TD algorithm similar to the real ones by rotating the direction of dummy trajectories. We evaluate their performance in three aspects: 1) *SE*, 2) *TE*, and 3) the time on DD_{degree}.

In our first simulation, we compare the number of dummy trajectories on *SE*. As shown in Fig. 4, *SE* is negatively correlated with the number of the dummy trajectories generated. With the increment of location exposure risk, which means that the degree of privacy protection is weaken, much dummy trajectory number decreases at the same time. On the contrary, the lower the location exposure risk, the more number of dummy trajectory generated, the higher the degree of privacy protection.

It can be obviously seen from Fig. 4 that under the same *SE* value, our RcDT scheme generates the most number of dummy trajectories than that of RPDT and TPPA-TD. This observation indicates that RcDT is more effective and can generate more dummy trajectories conforming to the requirements than the other two methods. The smaller the *SE* (the higher degree of the privacy protection), the larger the number of generated trajectories. The result shows that our RcDT scheme has obvious protection effects when the mobile users have high requirements on trajectory privacy.

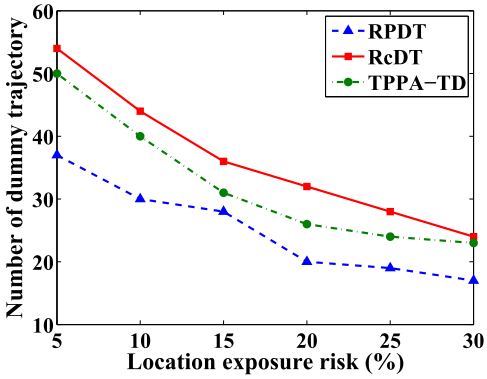


FIGURE 4. Comparison of the number of dummy trajectories on *SE*.

Then we compare the number of dummy trajectories on *TE*. As can be seen from Fig. 5, *TE* is negatively correlated with the number of the dummy trajectories generated as well. As the trajectory exposure risk increases, the number of dummy trajectories generated becomes smaller and smaller. The trajectory exposure risk is also negatively correlated with the degree of privacy protection. Therefore, when the trajectory exposure risk is higher, the degree of privacy protection is lower.

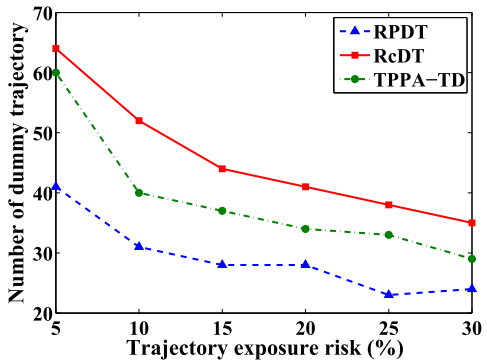


FIGURE 5. Comparison of the number of dummy trajectories on *TE*.

Fig. 5 shows that when the privacy protection requirements are higher, the number of dummy trajectories of RcDT is increasing. It is the second for RPDT and for TPPA-TD it

is the minimal under the best privacy protection degree. This observation indicates that our RcDT scheme is more suitable for the scenarios where users have higher requirements for privacy protection. When the value of trajectory exposure risk is constant, RcDT generates more dummy trajectories. That is, when the privacy protection effect of users is roughly equal, the probability that the real trajectories are recognized in RcDT is smallest, so the privacy protection effect of RcDT is the best.

In order to further evaluate the performance of RcDT, we compare the time on different distance deviation degree DD_{degree}, and its result is shown in Fig. 6. As we can see, the time required for TPPA-TD method is not much different from that of RcDT scheme. The results show that in the case of the same DD_{degree}, it takes more time for RPDT method to generate a dummy trajectory that meets the requirements than the RcDT method. When the mobile user does not require high degree of privacy protection, DD_{degree} can be larger. At this time, the time efficiency of RPDT and RcDT are similar, which means that RPDT can also meet the requirements. However, when users' privacy protection requirements are high, the time superiority of RcDT is more obvious.

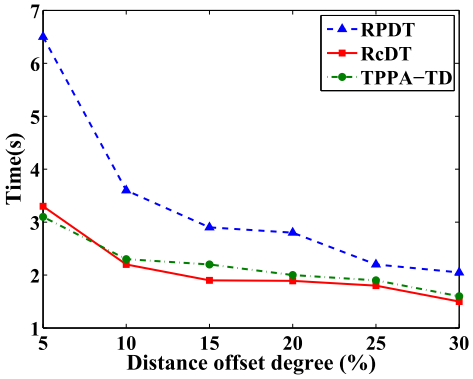


FIGURE 6. Comparison of time performance.

VII. CONCLUSION

Privacy protection is an important and challenging problem in MSNs. In this paper, we focus on the problem of trajectory privacy preservation in MSNs. We have proposed a privacy protection scheme based on R-constrained dummy trajectory (RcDT) which aims at the shortcomings of the traditional methods. The randomly generated dummy locations are range-constraint, making the similarity between the false trajectories and real trajectories enhanced. At the same time, considering the real location and trajectory exposure risks, the privacy degree of users' is improved. Finally, extensive experiments are carried out and the experimental results show that the *SE* and *TE* of our schema are superior than that of the state of the art benchmarks, and the execution time also fast under the same deviation degree, verifying that the proposed trajectory privacy protection scheme is more efficient and effective, and has some advantages over the existing methods.

The Euclidean distance is employed when calculating the distance in this paper. However, the actual situation may be more complicated. In many cases, the Euclidean distance may not be suitable. Therefore, we will consider other distance calculation methods in the future to study more practical trajectory privacy protection scheme. On the other hand, as discussed in this paper, the dummy locations are randomly generated around the real location within a radius, but there is a lack of the consideration that the generated dummy location could be some places that people cannot access to such as lakes, deserts or mountains etc. This proposal would be studied at our next stage of research.

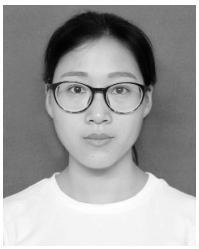
REFERENCES

[1] Y. Liang, Z. Cai, Q. Han, and Y. Li, "Location privacy leakage through sensory data," *Secur. Commu. Netw.*, vol. 2017, Dec. 2017, Art. no. 7576307.
[2] N. Capurso, T. Song, W. Cheng, J. Yu, and X. Cheng, "An Android-based mechanism for energy efficient localization depending on indoor/outdoor context," *IEEE Internet Things J.*, vol. 4, no. 2, pp. 299–307, Apr. 2017.
[3] Y. Liang, Z. Cai, J. Yu, Q. Han, and Y. Li, "Deep learning based inference of private information using embedded sensors in smart devices" *IEEE Netw. Mag.*, vol. 32, no. 4, pp. 8–14, Jul./Aug. 2018.
[4] T. Song, R. Li, B. Mei, J. Yu, X. Xing, and X. Cheng, "A privacy preserving communication protocol for IoT applications in smart homes," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 1844–1852, Dec. 2017.
[5] R. Li, T. Song, N. Capurso, J. Yu, J. Couture, and X. Cheng, "IoT applications on secure smart shopping system," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 1945–1954, Dec. 2017.
[6] L. Ni, C. Li, X. Wang, H. Jiang, and J. Yu, "DP-MCDBSCAN: Differential privacy preserving multi-core DBSCAN clustering for network user data," *IEEE Access*, vol. 6, pp. 21053–21063, May 2018.
[7] T. Song, N. Capurso, X. Cheng, J. Yu, B. Chen, and W. Zhao, "Enhancing GPS with lane-level navigation to facilitate highway driving," *IEEE Trans. Veh. Technol.*, vol. 66, no. 6, pp. 4579–4591, Jun. 2017.
[8] H. Chen and W. Lou, "On protecting end-to-end location privacy against local eavesdropper in wireless sensor networks," *Pervasive Mobile Comput.*, vol. 16, pp. 36–50, Jan. 2015.
[9] J. Zhang, Y. Yuan, X. Wang, L. Ni, J. Yu, and M. Zhang, "RPAR: Location privacy preserving via repartitioning anonymous region in mobile social network," *Secur. Commu. Netw.*, vol. 2018, Dec. 2018, Art. no. 6829326.
[10] X. Zheng, Z. Cai, J. Yu, C. Wang, and Y. Li, "Follow but no track: Privacy preserved profile publishing in cyber-physical social systems," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 1868–1878, Dec. 2017.
[11] S. Zhang, G. Wang, and Q. Liu, "A dual privacy preserving scheme in continuous location-based services," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 4191–4200, Oct. 2018.
[12] T. Wang, J. Zeng, M. Z. A. Bhuiyan, H. Tian, Y. Cai, Y. Chen, and B. Zhong, "Trajectory privacy preservation based on a fog structure for cloud location services," *IEEE Access*, vol. 5, pp. 7692–7701, 2017.
[13] X. Kong, M. Li, K. Ma, K. Tian, M. Wang, Z. Ning, and F. Xia, "Big trajectory data: A survey of applications and services," *IEEE Access*, vol. 6, pp. 58295–58306, 2018.
[14] X. Zheng, G. Luo, and Z. Cai, "A fair mechanism for private data publication in online social networks," *IEEE Trans. Netw. Sci. Eng.*, to be published. doi: 10.1109/tnse.2018.2801798.
[15] Z. He, Z. Cai, and J. Yu, "Latent-data privacy preserving with customized data utility for social network data," *IEEE Trans. Veh. Technol.*, vol. 67, no. 1, pp. 665–673, Jan. 2018.
[16] Q. Han, D. Lu, K. Zhang, X. Du, and M. Guizani, "Lclean: A plausible approach to individual trajectory data sanitization," *IEEE Access*, vol. 6, pp. 30110–30116, 2018.
[17] Z. Hu, J. Yang, and J. Zhang, "Trajectory privacy protection method based on the time interval divided," *Comput. Secur.*, vol. 77, pp. 488–499, Aug. 2018.
[18] A. Y. Ye, Y. Li, and L. Xu, "A novel location privacy-preserving scheme based on l-queries for continuous LBS," *Comput. Commun.*, vol. 98, pp. 1–10, Jan. 2017.

[19] K. Xing, C. Hu, J. Yu, X. Cheng, and F. Zhang, "Mutual privacy preserving *K*-means clustering in social participatory sensing," *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 2066–2076, Apr. 2017.
[20] Z. Huo, Y. Huang, and X. Meng, "History trajectory privacy-preserving through graph partition," in *Proc. 1st ACM Int. Workshop Mobile Location-Based Service (MLS)*, 2011, pp. 71–78.
[21] C.-Y. Chow and M. F. Mokbel, "Trajectory privacy in location-based services and data publication," *ACM SIGKDD Explorations Newslett.*, vol. 13, no. 1, pp. 19–29, 2011.
[22] M. Terrovitis and N. Mamoulis, "Privacy preservation in the publication of trajectories," in *Proc. 9th IEEE Int. Conf. Mobile Data Manage. (MDM)*, Apr. 2008, pp. 65–72.
[23] L. Huang, K. Matsuura, H. Yamane, and K. Sezaki, "Enhancing wireless location privacy using silent period," in *Proc. IEEE Wireless Commun. Netw. Conf.*, vol. 2, Mar. 2005, pp. 1187–1192.
[24] R.-H. Hwang, Y.-L. Hsueh, and H.-W. Chung, "A novel time-obfuscated algorithm for trajectory privacy protection," *IEEE Trans. Services Comput.*, vol. 7, no. 2, pp. 126–139, Apr./Jun. 2014.
[25] J. Zhao, Y. Zhang, X. Li, and J. Ma, "Trajectory privacy protection method based on trajectory frequency suppression," *Chin. J. Comput.*, vol. 37, no. 10, pp. 2096–2106, 2014.
[26] Z. Huo, X. Meng, and Y. Huang, "Privatecheckin: A trajectory privacy protection method in mobile social networks," *J. Comput.*, vol. 36, no. 4, pp. 716–726, 2013.
[27] D. Yang, D. Zhang, B. Qu, and P. Cudré-Mauroux, "PrivCheck: Privacy-preserving check-in data publishing for personalized location based services," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput. (PUC)*, 2016, pp. 545–556.
[28] A. Hossain, A. Quattrone, E. Tanin, E. Tanin, and L. Kulik, "On the effectiveness of removing location information from trajectory data for preserving location privacy," in *Proc. 9th ACM SIGSPATIAL Int. Workshop Comput. Transp. Sci. (CTS)*, 2016, pp. 49–54.
[29] T.-H. You, W.-C. Peng, and W.-C. Lee, "Protecting moving trajectories with dummies," in *Proc. IEEE Int. Conf. Mobile Data Manage. (MDM)*, May 2007, pp. 278–282.
[30] Z. Huo, X. Meng, H. Hu, and Y. Huang, "You can walk alone: Trajectory privacy-preserving through significant stays protection," in *Proc. Int. Conf. Database Syst. Adv. Appl. (DSAA)*, 2012, pp. 351–366.
[31] L. Ni, Y. Yuan, X. Wang, J. Yu, and J. Zhang, "A privacy preserving algorithm based on R-constrained dummy trajectory in mobile social network," *Procedia Comput. Sci.*, vol. 129, pp. 420–425, Jan. 2018.
[32] T. Peng, Q. Liu, D. Meng, and G. Wang, "Collaborative trajectory privacy preserving scheme in location-based services," *Inf. Sci.*, vol. 387, pp. 165–179, May 2017.
[33] R. Yarovoy, F. Bonchi, and L. Lakshmanan, "Anonymizing moving objects: How to hide a mob in a crowd?" in *Proc. ACM 12th Int. Conf. Extending Database Technol. (EDT)*, 2009, pp. 72–83.
[34] P.-R. Lei, W.-C. Peng, I.-J. Su, and C.-P. Chang, "Dummy-based schemes for protecting movement trajectories," *J. Inf. Sci. Eng.*, vol. 28, no. 2, pp. 335–350, 2012.
[35] S. Gao, J. Ma, W. Shi, and G. Zhan, "LTPPM: A location and trajectory privacy protection mechanism in participatory sensing," *Wireless Commun. Mobile Comput.*, vol. 15, no. 1, pp. 155–169, 2015.
[36] *Thomas Brinkhoff Network-Based Generator Moving Objects*. Accessed: Apr. 21, 2019. [Online]. Available: http://iapg.jade-hs.de/personen/brinkhoff/generator/

JINQUAN ZHANG received the Ph.D. degree in computer science and technology from Tongji University, Shanghai, China, in 2007. He is currently an Associate Professor with the College of Computer Science and Engineering, Shandong University of Science and Technology, Shandong. His current areas of research are cloud computing, Petri net, privacy preservation, machine learning, and parallel and distributed processing. He is a Senior Member of the China Computer Federation (CCF).

XIAO WANG is currently pursuing the M.S. degree with the College of Computer Science and Engineering, Shandong University of Science and Technology. Her main research interests include privacy preservation, cloud computing, and machine learning.



YANFENG YUAN is currently pursuing the M.S. degree with the College of Computer Science and Engineering, Shandong University of Science and Technology. Her main research interests include privacy preservation, cloud computing, and machine learning.

LINA NI received the Ph.D. degree in computer software and theory from Tongji University, Shanghai, China, in 2009. She is currently an Associate Professor with the College of Computer Science and Engineering, Shandong University of Science and Technology, Shandong. She is also with the Key Laboratory of the Ministry of Education for Embedded System and Service Computing, Tongji University. Her current areas of research are cloud computing, privacy preservation, Petri net, distributed algorithms, machine learning, and intelligent computing. She is a member of the ACM and a Senior Member of the China Computer Federation (CCF). She is also a Committee Member of the Professional Committee of Network Information Service of China Automation Federation.