IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 25, 2017, accepted April 11, 2017, date of publication May 3, 2017, date of current version June 7, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2698078

Trajectory Privacy Preservation Based on a Fog Structure for Cloud Location Services

TIAN WANG[1], JIANDIAN ZENG[1], MD ZAKIRUL ALAM BHUIYAN[2], (Member, IEEE), HUI TIAN[1], YIQIAO CAI[1], YONGHONG CHEN[1], AND BINENG ZHONG[1]

[1]Department of Computer Science and Technology, Huaqiao University, Xiamen, 361021, China
[2]Department of Computer and Information Sciences, Fordham University, Bronx, NY 10458, USA

Corresponding author: Hui Tian (cshtian@126.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61672441 and in part by the Foster Project for Graduate Student in Research and Innovation of Huaqiao University under Grant 1511414002.

ABSTRACT The development of mobile cloud computing technology has made location-based service (LBS) increasingly more popular. Given the continuous requests to cloud LBS servers, the amounts of location and trajectory information collected by LBS servers are continuously increasing. Privacy awareness for LBS has been extensively studied in recent years. Among the privacy concerns about LBS, trajectory privacy preservation is particularly important. Based on privacy preservation models, previous work have mainly focused on peer-to-peer and centralized architectures. However, the burden on users is heavy in peer-to-peer architectures, because user devices need to communicate with LBS servers directly. In centralized architectures, a trusted third party (TTP) is introduced, and acts as a bridge between users and the LBS server. Anonymity technologies, such as k-anonymity, mix-zone, and dummy technologies, are usually implemented by the TTP to ensure safety. There are certain drawbacks in TTP architectures: Users have no physical control of the TTP. Moreover, the TTP is more attractive to adversaries, because substantially more sensitive information is stored by the TTP. To solve the above-mentioned problems, in this paper, we propose a fog structure to store partial important data with the dummy anonymity technology to ensure physical control, which can be considered as absolutely trust. Compared with cloud computing, fog computing is a promising technique that extends the cloud computing to the edge of a network. Moreover, fog computing provides local computation and storage abilities, wide geo-distribution, and support for mobility. Therefore, mobile users' partial important information can be stored on a fog server to ensure better management. We take the principles of similarity, intersection, practicability, and correlation into consideration and design a dummy rotation algorithm with several properties. The effectiveness of the proposed method is validated through extensive simulations, which show that the proposed method can provide enhanced privacy preservation.

INDEX TERMS LBS privacy, trajectory privacy preservation, fog computing, rotation.

I. INTRODUCTION

Recent years have brought significant growth in the number of location-aware devices with the development of location sensing technologies, including smart phones, RFID and wireless sensors [1], [2]. The popularity of mobile devices in people's daily life has motivated a series of applications [3], [4], especially in energy-efficient mobile cloud computing. It is commonly believed that when a person comes to an unfamiliar place, he/she wants to know if there is a supermarket or hotel nearby. Location based services (LBS) are used to address this issue. Usually, mobile users continuously send queries to cloud LBS servers [5]. If we order this information according to sequences of time and positions, users' trajectory information can be obtained, which may pose a threat to these users if such data is leaked. For example, concerning check-in applications [6], many shops encourage users to check-in, and provide points to users for discounts. With increasing number of check-ins, a greater privacy disclosure probability results. With large numbers of check-ins, isolated check-in positions can be used to form a trajectory over time. As shown in Fig. 1, the black points denote the check-in positions. If we connect these black points, a trajectory is formed, which may expose a user's personal information about where or when they travel.

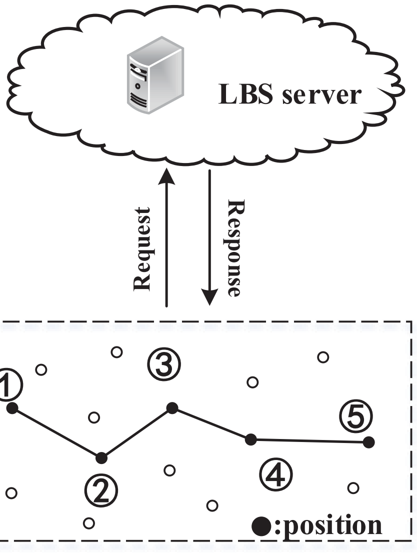


FIGURE 1. Traditional LBS system model.

Therefore, trajectory privacy preservation is a critical issue when using location-based services.

Typical traditional LBS system models are also called peer-to-peer architectures. Users first send query requests to a cloud LBS server. Then, the cloud LBS server provides information services to users according to their demands. However, once the service data is captured by adversaries, users' position privacy is compromised, which may lead to the inference of sensitive information about individuals. Moreover, the energy consumption is excessive because users' devices need to communicate with the LBS server directly. To alleviate this burden, a trusted third party (TTP) is introduced into the system, thereby forming a centralized architecture [7], [8]. In centralized architectures, the TTP acts as a bridge between users and the LBS server: Users first send query request to TTP. Encryption or anonymity methods, such as k-anonymity [9], mix-zone [10], and dummy technologies [11], are implemented by the TTP to ensure safety. The TTP uploads the processed data to the LBS server [12]. Finally, the LBS server returns information services to the TTP and the users. Gedik and Liu [13] used a trusted anonymity server to perform spatio-temporal cloaking, which allows users to define and modify their location privacy specifications at the granularity of single messages. However, there are some drawbacks to TTP architectures: Users have no physical control of the TTP. Moreover, the TTP is more attractive to adversaries because substantially more sensitive information is stored by the TTP [14].

To solve the above-mentioned problems, a fog server can be introduced [15]. Fog computing is closer to local and was first proposed by Cisco in 2011 [16]. Fog computing provides local computation and storage abilities, wide geo-distribution and supports for mobility [17], [18]. Compared to cloud computing, fog computing is a promising technique that extends cloud computing to the edge of the network, thus enabling new applications and services. Mobile users' partial important information can be stored in the fog server to ensure physical control; therefore, fog servers can be considered as absolutely trusted [19]. In this paper, we use a fog server to replace a traditional TTP server (the fog-based architecture is introduced in Section III) and propose a dummy rotation algorithm (DR) to ensure the anonymity. The main contributions are listed as follows:

1. We use a fog server to store partial information to ensure physical control.

2. We propose a DR algorithm to hide the real trajectory among dummy trajectories, whereby the principles of similarity, intersection, practicability and correlation are considered. Therefore, there are two forms of insurances in our method.

3. We conduct extensive simulations, and the effectiveness of our proposed solution is validated.

The remainder of this paper is organized as follows. Section II reviews related work. Section III presents the system model and corresponding preliminaries. Section IV introduces the proposed dummy algorithm. Security analyses are discussed in Section V. Simulation results are demonstrated in Section VI. Section VII concludes the paper.

II. RELATED WORK

LBS provides users with various services based on their geographical locations. In LBS, mobile users need to upload their locations, obtained from GPS, GSM or CDMA, to the LBS server for accessing data services [20], [21]. In this way, real-time location information is stored by the LBS server. However, sensitive information may be leaked because the LBS server is not absolutely trusted.

Privacy awareness in LBS has been extensively studied in recent years. Among the privacy concerns about LBS, trajectory privacy preservation is particularly important because the trajectory information may expose a user's personal information about where or when they travel. In addition to basic cryptography protection schemes, the majority of trajectory privacy preservation methods can be divided into three techniques: The first is spatial cloaking [22], which is a method for hiding a user's real location in a cloaked area. For instance, when k users that are in the same area submit queries to the LBS server, the k queries are packaged as one query to request service. Therefore, their real locations are replaced by the area composed of these k users, which somehow protects a user's location privacy. However, the k-anonymity technology is vulnerable to correction attacks [9]. Based on k-anonymity technique, Hwang *et al.* [23] proposed a novel time-obfuscated technique to obtain a set of similar trajectories by breaking the sequence of the query issuing time. The mechanism finds (r-1) trajectories to be distributed with a distance variance and maintains a relatively low indexing cost. The second technique is mix-zones, which is a method that changes pseudonyms when a set of users are added to mix zones [24]. When a user enters a mix zone, a new, unused pseudonym is used. Ying *et al.* [10] proposed a dynamic mix zone for vehicular user that can support different privacy levels and ignore a user's location at the time of request. Compared to traditional fixed mix-zone techniques, the proposed scheme can be used for safety applications on certain types of roads. However, users' location information cannot be updated until existing the mix zone [25], which can cause delayed services. The third technique is dummy trajectories [26]. In this technology, a user's real trajectory is used as a generator to obtain multiple dummy trajectories to confuse attackers. Tang *et al.* [11] focused on long-term location privacy protection, in which symmetry, decongestion, practicability, and consistency were proposed for dummy trajectory generation. They proposed and developed a set of novel dummy generation algorithms, therein considering both real geographical information and long-term consistency. However, the burden on the LBS server will be higher because multiple dummy trajectories along with a real trajectory are uploaded to the LBS server simultaneously. This is acceptable because it is necessary to generate multiple trajectories to achieve anonymity. Moreover, the other two methods upload k queries to the LBS server simultaneously as well.

Based on the trajectory privacy preservation model, peer-to-peer and centralized architectures have been developed. In peer-to-peer architectures [27], mobile users upload queries to the LBS server directly. Users can also communicate with other peers through a wireless network. Che *et al.* [28] presented a dual-active spatial cloaking algorithm for mobile peer-to-peer networks. The proposed algorithm enables peers to achieve required anonymity goal in less time per query and uses peer location information and location records. The drawback of this architecture is that the overhead to users is larger. In centralized architectures [13], [29], a TTP is introduced, therein acting as an intermediate tier between users and the LBS server [9]. Gao *et al.* [29] proposed a personalized anonymization model to select k-anonymity trajectory, in which trajectory preprocessing, an optimal trajectory graph model and an anonymization set section are implemented in the proposed model. In the above centralized architecture, privacy preservation algorithms are usually implemented by the TTP to achieve anonymity. However, there are several challenges to centralized architectures: 1) Users have no physical control of the TTP. 2) The TTP is more attractive to adversaries because substantially more sensitive information is stored in the TTP. 3) The social pattern and geographic map relationships should be considered by the TTP [20].

Compared to the TTP structure, a fog server is closer to local [30]. The concept of fog computing was first proposed by Cisco. Fog computing is characterized by local computation and storage abilities, wide geo-distribution and support for mobility [17], [18]. In regard to data storage and management, a fruitful interplay between the Cloud and the Fog exists. The main differences between cloud and fog computing is that fog computing can provide local storage and mobility for end users. Mobile users' partial important information can be stored in a fog server to ensure physical control [31]. Based on the local computations in the fog server, the social pattern and geographic map relationship can be considered. In this paper, we introduce a fog sever between the mobile users and the LBS server, whereby the fog server is closer to users. Certification, anonymity and encryption are implemented in the fog server. Moreover, the dummy trajectory technology is adopted to achieve anonymity.

III. PRELIMINARIES

In this section, we first present the system architecture of the proposed method and then introduce the attack model. Finally, we present privacy metrics.

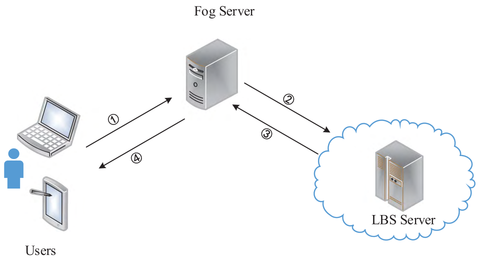


FIGURE 2. Illustration of the proposed system architecture.

A. SYSTEM ARCHITECTURE

As shown in Fig. 2, there are three parts to the proposed system architecture: users, the fog server, and the LBS server, respectively. A fog server is implemented in user's spare machines that have sufficient hard drive space. The fog server is between the users and the LBS server and is owned by users. The main workflow can be described as follows:

① The location information is uploaded to fog server with a corresponding signature; the query can be denoted as (UserID, query time, location).

② The fog server is responsible for generating dummy trajectories based on previous locations. To ensure data security, an encryption algorithm, such as RSA, can be implemented. Then, the fog server stores the partial data and sends other encrypted data to the LBS server for further requests.

③ The LBS server decrypts the data and obtains corresponding query services. Then, the LBS server sends the query services to the fog server.

④ The fog server receives the information services from the LBS server, and then, it sends to the users based on corresponding UserIDs.

B. ATTACK MODEL

Given the role of an attacker, attacks can be classified as either insider threats and external attacks.

With insider threats, a system manager may leak information on purpose. Moreover, a user may share position information on social platforms (e.g., by uploading photos to Facebook or Twitter). In this way, a user's partial positions are exposed to attackers.

With external attacks, there are active attacks and passive attacks. In active attacks, attackers send malicious information to mislead users, such as in luring attacks [32]. In passive attacks, attackers collect geographic and social information to estimate a user's real location information [33].

Based on the analyses above, we assume that attacks can obtain a user's partial real positions or all information from the LBS server. In addition, we assume that the fog server is trusted unless users lose physical control over it.

C. PRIVACY METRIC

To measure the performance of privacy preservation, there are several metrics used by our system.

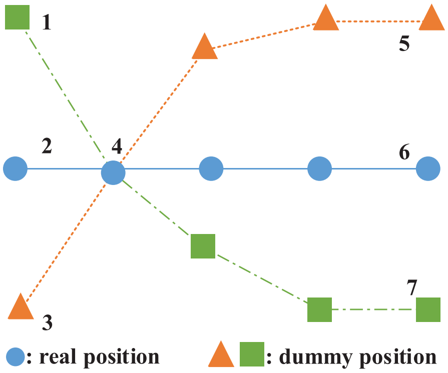


FIGURE 3. Illustration of TDP.

1) TRAJECTORY DISCLOSURE PROBABILITY (*TDP*)

Suppose that there are several intersections among dummy trajectories and the real trajectory. More intersections correspond to a lower *TDP*. For example, in Fig. 3, the real trajectory is (2, 4, 5), and there is only one intersection among the trajectories. In addition to dummy trajectories (1, 4, 7) and (3, 4, 5), there are six extra synthetic trajectories (i.e., (1, 4, 5) and (1, 4, 7)). In this case, the trajectory disclosure probability is 1/9. Thus, in this paper, the disclosure probability can be denoted as $TDP = 1/N_t$, where N_t is the number of all trajectories.

2) POSITION DISCLOSURE PROBABILITY (*PDP*)

The position disclosure probability can be determined as $PDP = k/N_p$, where k and N_p denote the number of real positions and all positions, respectively. It can be seen that when the number of dummy trajectories increases, N_p increases, and *PDP* decreases.

3) AVERAGE EUCLIDEAN DISTANCE (*AED*)

In our method, we generate multiple dummy trajectories based on the real trajectory. If the distance is too small or too large, the privacy level will be lower. Thus, the distance between dummy trajectories and the real distance should be considered. In this way, we set the average Euclidean distance as a metric, which can be calculated as follow:

$$\overline{AED} = \frac{1}{n} \times \sum_{i=1}^{n} dis_{i,r} \tag{1}$$

where $dis_{i,r}$ denotes the Euclidean distance between the i^{th} dummy trajectory and real distance.

TABLE 1. Notations used in Section IV.

Notation	Implications
Δ	Step of rotation angle increment
θ	Rotation angle
k	Number of dummy trajectories
T_r	Real trajectory
D_i	Dummy trajectory
T_c	Current trajectory
AGD	Average Gaussian distance
AED	Average Euclidean distance

4) LOCAL DATA VOLUME (*LDV*)

As partial data is stored in fog server to ensure physical control, the data volume in fog server should be considered. If data volume is too small, the raw data may be restored by attackers. The related technologies are mean imputation [34], regression imputation [35], multiple imputation [36] and so on. In the other hand, it is not necessary to store too much data in fog serve because the burden of fog server will be huge. So, the local data volume should be suitable in terms of security.

IV. DUMMY ROTATION ALGORITHM

Before introducing the DR algorithm, several principles should be addressed.

A. SIMILARITY

The dummy trajectories should be similar to the real trajectory, which provide a better ability to mislead attackers.

B. INTERSECTION

The dummy trajectories should have some intersections with the real trajectory to confuse attackers. However, more intersections means a higher *PDP* and a smaller *TDP*. Thus, there should be a balance between *PDP* and *TDP*.

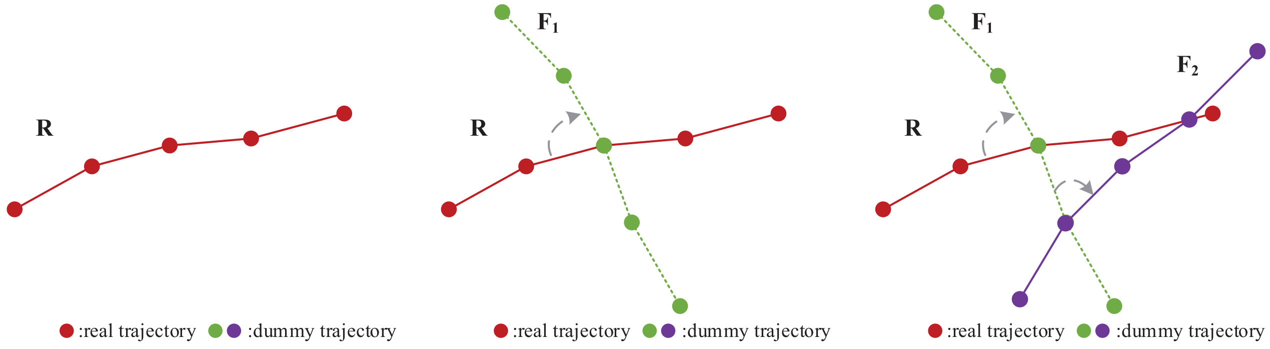
C. PRACTICABILITY

Some positions in dummy trajectories may be impractical, such as on mountains or rivers. So, the geographic map should be considered, and the impractical locations should be replaced.

D. CORRELATION

The dummy trajectories should have correlations with the user's location history. Therefore, social pattern should be considered, and the dummy trajectories should be in the range of the user's social circle.

Based on the above principles, we proposed a dummy rotation algorithm, that chooses an appropriate angle to rotate. If the rotation angle is too small or too large, the dummy trajectories will be too close to the real trajectory, which may result in a lower privacy level. Therefore, the rotation angle is closely related to the trajectory privacy level. Moreover, the



(a) (b) (c)

FIGURE 4. Illustration of dummy trajectory generation: (a) Real trajectory. (b) Single rotation. (c) Two rotations.

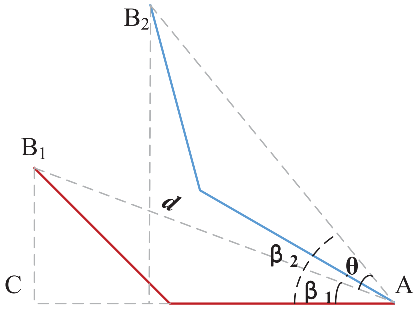


FIGURE 5. Illustration of position calculation.

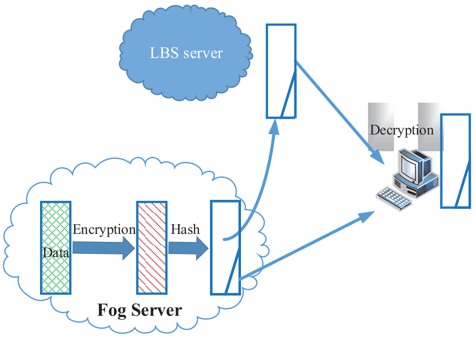


FIGURE 6. Illustration of local storage.

average Euclidean distance between dummy trajectories and the real trajectory varies with the increment of the rotation angle. In this way, we can conclude that the trajectory privacy and rotation angle follow a Gaussian distribution. To find a suitable rotation angle, we first generate samples by increasing the rotation angle gradually. The corresponding *AED* can be obtained by formula (1). Then, the i^{th} average Gaussian distance (AGD_i) can be calculated as follow:

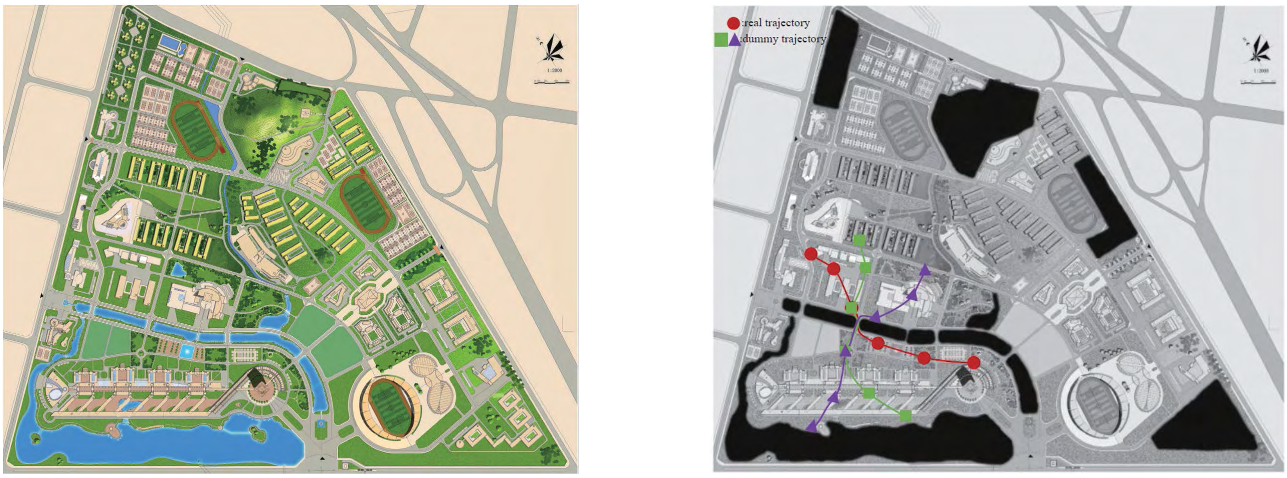
$$AGD_i = \frac{1}{\sqrt{2\pi}\sigma} \times e^{-\frac{(\overline{ADE}-dis_{i,r})^2}{2\sigma^2}} \qquad (2)$$

Based on formula (2), the maximum *AGD* and the corresponding rotation angle can be obtained. The pseudo-code and related notations are shown in Algorithm 1 and Table 1, respectively.

```
Algorithm 1 Obtain Rotation Angle
Input: real trajectory ($T_r$), step of rotation angle increment ($\Delta$), number of dummy trajectories ($k$)
Output: final rotation angle ($\theta_s$)
for $\theta = \Delta : \Delta : \frac{2\pi}{k}$ do
    for i=1:1:k do //generate $k$ dummy trajectories
        $d_i = getDummy(T_r, \theta, k)$; // get dummy trajectory
    end for
end for
Get $AED$ and $AGD$ by formula (1) and (2);
$\theta_s$ = corresponding $\theta$ with maximum $AGD$;
```

```
Algorithm 2 getDummy($T_r, \theta, k$) (Get Dummy Trajectories)
Input: real trajectory ($T_r$), rotation angle ($\theta_s$), number of dummy trajectories ($k$)
Output: dummy trajectories
$T_c = T_r$; //set $T_r$ as the current trajectory
for i =1:1:k do
    select a rotation point in $T_c$ randomly and then rotate clockwise by $\theta_s$ to generate a new trajectory;
    check all positions and replace impractical positions to obtain $d_i$;
    $T_c = T_i$ //set $T_i$ as the current trajectory
end for
```

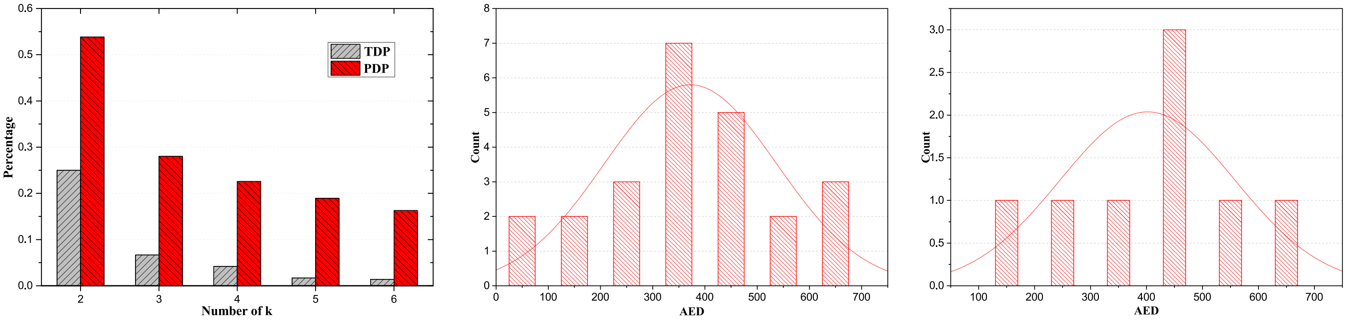
After obtaining a suitable rotation angle using algorithm 1, we generate dummy trajectories using algorithm 2. To obtain multiple dummy trajectories, as shown in Fig. 4, we randomly select a rotation point along the real trajectory and then rotate clockwise by θ_s to generate a new trajectory. Finally, we check for impractical positions and replace them by randomly



(a)

(b)

FIGURE 7. (a) The satellite map of a university. (b)The satellite map with marks.



(a)

(b)

(c)

FIGURE 8. Privacy level and Gaussian distribution: (a) *TDP* and *PDP* vs. *k*. (b) *AED* Gaussian distribution when $\Delta = 5$. (c) *AED* Gaussian distribution when $\Delta = 15$.

selecting a position from an impractical position's neighbor (Algorithm 2, lines 3-4).

The detailed method of calculating dummy trajectories is illustrated in Fig. 5. In this figure, the real trajectory is denoted as ACB_1, and the rotation angle $\angle B_1AB_2$ is θ. $\angle CAB_1$ can be obtained as follows:

$$\beta_1 = arctan\frac{B_1.y - A.y}{B_1.x - A.x}; \tag{3}$$

Then, $\beta_2 = \beta_1 + \theta$. Therefore, the position of B_2 is calculated as follow:

$$\begin{cases} B_2.x = A.x - d * cos(\beta_2) \\ B_2.y = A.y - d * sin(\beta_2) \end{cases} \tag{4}$$

Theorem 1: The time complexity of the rotation algorithm is $O(m*n)$, where m and n denote the number of samples and dummy trajectories, respectively.

Proof: In the proposed algorithm, we first generate multiple trajectory samples (m) by incrementing the rotation angle (Algorithm 1 lines 1-5), in which the loop is performed m times. Then, in each trajectory sample, we obtain n dummy trajectories (see Algorithm 2), whose time complexity is $O(n)$. Therefore, the total time complexity is $O(m*n)$. □

Theorem 2: The rotation angle θ should be in $[0, \frac{2\pi}{k}]$, where k is the number of dummy trajectories.

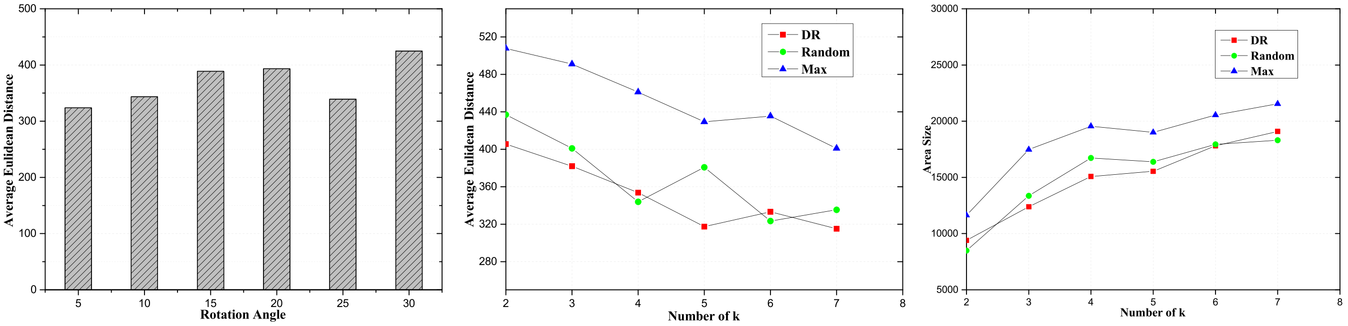
Proof: In our method, the dummy trajectories are generated by rotating the current trajectory by θ. The angle of one circle is 2π, and the number of dummy trajectories is k. If the rotation angle θ is larger than $\frac{2\pi}{k}$, there will be some redundant samples because all samples are generated by incrementing with a step angle Δ. For example, if the max $\theta = \Delta + \frac{2\pi}{k}$, the samples that are generated with Δ will be redundant. Therefore, the rotation angle θ should be in $[0, \frac{2\pi}{k}]$. □

V. SECURITY ANALYSES

In this section, we discuss the security of our proposed method.

Case 1: Partial information from the LBS server is leaked.

In this case, an attacker can obtain partial information from the LBS server, and the data on the fog server are safe.



(a)

(b)

(c)

FIGURE 9. ***AED*** **variation: (a)** ***AED*** **vs. rotation angle. (b)** ***AED*** **vs.** ***k*****. (c) Area size vs.** ***k*****.**

In our method (as shown in Fig. 6), the data is encrypted by asymmetric algorithms (such as RSA and Elgamal) before uploading the data to the LBS server. Then, the encrypted data is divided into two parts. The partial data is stored on the fog server to ensure physical control, and other data is uploaded to the LBS server. In this way, even if the data on the LBS server are leaked, attacks still cannot restore the raw data because of lacking data that are stored on the fog server.

Case 2: The LBS server is hijacked, and the attacker has obtained data from both the fog server and the LBS server.

In this case, the attacker hijacks the LBS server, and sends malicious requests to the fog server. Suppose that the attacker has obtained the data from both the fog server and the LBS server. In this way, even if the attacker has decrypted the data successfully and obtained all the trajectories, it is still difficult for the attacker to distinguish the real trajectory because of the DR algorithm applied with the four above-mentioned principles. For example, to prevent typical correlation attacks [9], the generated dummy trajectories are in the range of the user's social pattern. In addition, we replace impractical positions to make the dummy trajectories more realistic. Specifically, there are two types of insurances in our method.

Case 3: A user's partial real positions are leaked.

In the worst case, assume that all trajectories and partial real positions are accessed by the attacker. In this case, there is some probability for the attacker to obtain the real trajectories. Because there are intersections among the generated trajectories, if the leaked real positions are exactly the intersections, the attacker still cannot distinguish the real trajectory. However, if the real positions are not all intersections, the attacker may obtain the real trajectory. However, the conditions are critical because it is difficult to obtain all the trajectories of a user if a strong encryption algorithm is applied.

VI. PERFORMANCE EVALUATION

In this section, extensive simulations are conducted using Matlab 2015a to validate the performance of our proposed solution. To simulate actual scene, we use a satellite map of our university (see Fig. 7(a)) and collect a student's daily movements as samples. Fig. 7(b) indicates the corresponding impractical position in black; real positions are shown in red, and dummy positions are shown in green and purple.

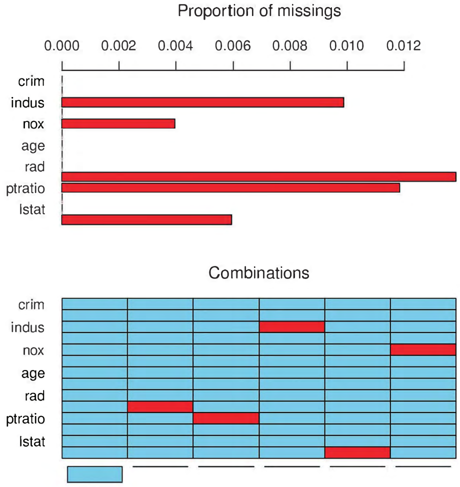


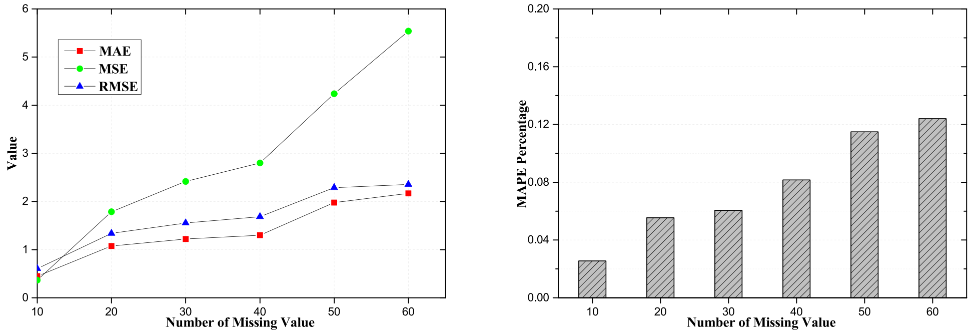
FIGURE 10. **The proportions and combinations of missing values.**

For comparisons, three different types of solutions for rotation are demonstrated. The first type is our proposed solution *DR*. The second type is *Random*. As the name suggests, the rotation angle is randomly generated in $[0, \frac{2\pi}{k}]$. The third type is *Max*, which rotates by $\frac{2\pi}{k}$ to generate dummy trajectories.

A. PRIVACY LEVEL AND GAUSSIAN DISTRIBUTION

Fig. 8(a) shows the trends of *TDP* and *PDP* when the number of dummy trajectories increases from 2 to 6, in which $\Delta = 5$. It can be observed that both *TDP* and *PDP* exhibit a decreasing trend with increasing k. Because the total numbers of trajectories and positions increase with increasing k and the numbers of real trajectories and positions are constant, *TDP* and *PDP* decrease.

In Figs. 8(b) and 8(c), Gaussian distributions of the average Euclidean distance are presented when $\Delta = 5$ and $\Delta = 15$, respectively. It can be observed that the number of *AED* is greater when *AED* is close to $\overline{AED}$. A smaller Δ, results in



(a) (b)

FIGURE 11. **Data recovery with mean imputation: (a) *MAE*, *MSE*, *RMSE* vs. Number of missing values. (b) *MAPE* vs. Number of missing values.**

larger samples. Therefore, a better the rotation angle can be obtained with smaller Δ.

From Fig. 8, we can conclude that *TDP* and *PDP* are lower under greater numbers of dummy trajectories. Moreover, with smaller Δ, the Gaussian distributions of *AED* are more reasonable, and a suitable rotation angle can be obtained.

B. AED VARIATION SIMULATIONS

In this section, we consider the influence of the rotation angle θ and different dummy trajectories k. Fig. 9(a) shows *AED* variations when θ increases from 5 to 30, in which the step angle (Δ) is 5 and $k = 2$. This figure shows that the trend characterizing the *AED* of our proposed method is increasing when θ is in [5, 20]. However, there is a fluctuation when θ is greater than 25. That is because we randomly select a rotation point in the current trajectory. When θ is small, the differences are not obvious; however, when θ increases, the randomness cannot be ignored.

The variations in *AED* are shown in Fig. 9(b) when k increases from 2 to 7 and $\Delta = 5$. It can be observed that the overall trends of the three methods are decreasing trends. The *AED* in Random is fluctuating because the rotation angle is randomly selected. When $k = 6$, there is an increment of both Max and DR. That is because a new dummy trajectory is generated by rotating θ from the current trajectory. When the number of dummy trajectories k is sufficiently large, the deviation from the real trajectories increases.

Fig. 9(c) shows the size of the anonymity area under the three methods when k increases from 2 to 7 and $\Delta = 5$. In our paper, the size of anonymity area is calculated as $(x_{max} - x_{min}) * (y_{max} - y_{min})$, where x_{max} (y_{max}) and x_{min} (y_{min}) are the maximum and minimum coordinates, respectively. As shown in the figure, with the increasing k, the overall trends of these three methods are increasing trends. When the number of dummy trajectories k is greater than 5, there is a slight increment of the size of the anonymity area under the three methods. This is because when k increases, the dummy trajectories are almost throughout the entire map.

In Figs. 9(a), 9(b) and 9(c), it can be observed that there is a fluctuation when the rotation angle or k is bigger than a certain value. This has been explained in the above analyses. Another reason is that some impractical positions are replaced when generating dummy trajectories, and there exists randomness.

C. DATA RECOVERY SIMULATIONS

In this section, experiments concerning data recovery are conducted. Because partial data is stored on the fog server, we need to know how much data should be stored on the local fog server such that these data can not be restored given the remaining data. In this paper, mean imputation is used to recover the missing data. The main idea behind mean imputation is to replace any missing value with the mean of that variable for all other cases, which has the benefit of not changing the sample mean for that variable. The R (version 3.1.1) programming language is used for the experiments. We choose the BostonHousing dataset, which includes 14 variables and 504 values. Because there are no null values in BostonHousing dataset, we randomly select 25 values to be null as missing values. The proportions and combinations of missing values are shown in Fig. 10. It can be observed that the null values are mainly centralized in five variables.

To measure the performance of data recovery, the mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE) are used as metrics. Fig. 11(a) shows that the MAE, MSE and RMSE are influenced by the variations in the missing values. It can be observed that the overall trends of the three metrics are increasing trends with increasing numbers of missing values. When the number of missing values is greater than 40 (approximately 8% of all values), the errors given by the three metrics increase obviously. This is because when the number of missing values is sufficiently large, the mean values are strongly influenced. Therefore, the errors increase.

In Fig. 11(b), the variations in the MAPE are shown when the number of missing values increases from 10 to 60. The MAPE is a measure of the prediction accuracy of a forecasting method in statistics. The figure shows that the MAPE increases with increasing number of missing values. When there are approximately 10% missing values, the MAPE is greater than 10%. Thus, we can store 10% or more of the data on the fog server to prevent restoration of the data.

VII. CONCLUSION

With the development of location sensing technologies, the number of location-aware devices has increased rapidly in recent years. Among the concerns about LBS privacy preservation, trajectory privacy preservation is a critical topic. In this paper, we propose to use a fog server to store partial important data that can be physical controlled by users. A dummy rotation algorithm is designed considering the principles of similarity, intersection, practicability and correlation. Specifically, there are two types of insurances in the proposed fog structure. The simulation results show that the proposed method can achieve enhanced privacy preservation. Moreover, approximately 10% to 15% of the data can be stored on the local fog server to prevent restoration of the data by attackers.

REFERENCES

[1] T. Wang *et al.*, "Following targets for mobile tracking in wireless sensor networks," *ACM Trans. Sensor Netw.*, vol. 12, no. 4, p. 31, 2016.

[2] Z. Xia, X. Wang, L. Zhang, X. Sun, K. Ren, and Z. Qin, "A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 11, pp. 2594–2608, Nov. 2016.

[3] Z. Xia, X. Wang, X. Sun, and Q. Wang, "A secure and dynamic multi-keyword ranked search scheme over encrypted cloud data," *IEEE Trans. Parallel Distrib. Syst.*, vol. 27, no. 2, pp. 340–352, Jan. 2016.

[4] G. Han, J. Jiang, M. Guizani, and J. J. P. C. Rodrigues, "Green routing protocols for wireless multimedia sensor networks," *IEEE Wireless Commun.*, vol. 23, no. 6, pp. 140–146, Dec. 2016.

[5] Z. Fu, X. Sun, S. Ji, and G. Xie, "Towards efficient content-aware search over encrypted outsourced data in cloud," in *Proc. IEEE INFOCOM Conf. Comput. Commun.*, Apr. 2016, pp. 1–9.

[6] D. Yang, D. Zhang, B. Qu, and P. C. Mauroux, "Privcheck: Privacy-preserving check-in data publishing for personalized location based services," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2016, pp. 545–556.

[7] X. Li, M. Miao, H. Liu, J. Ma, and K. C. Li, "An incentive mechanism for k-anonymity in LBS privacy protection based on credit mechanism," in *Proc. Soft Comput.*, 2016, pp. 1–11.

[8] G. Jia, G. Han, J. Jiang, and L. Liu, "Dynamic adaptive replacement policy in shared last-level cache of dram/pcm hybrid memory for big data storage," *IEEE Trans. Ind. Informat.*, to be published, doi: 10.1109/TII.2016.2645941.

[9] 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.

[10] B. Ying, D. Makrakis, and H. T. Mouftah, "Dynamic mix-zone for location privacy in vehicular networks," *IEEE Commun. Lett.*, vol. 17, no. 8, pp. 1524–1527, Aug. 2013.

[11] F. Tang, J. Li, I. You, and M. Guo, "Long-term location privacy protection for location-based services in mobile cloud computing," *Soft Comput.*, vol. 20, no. 5, pp. 1735–1747, 2016.

[12] Z. Fu, F. Huang, X. Sun, A. Vasilakos, and C. N. Yang, "Enabling semantic search based on conceptual graphs over encrypted outsourced data," *IEEE Trans. Serv. Comput.*, to be published, doi: 10.1109/TSC.2016.2622697.

[13] B. Gedik and L. Liu, "Protecting location privacy with personalized k-anonymity: Architecture and algorithms," *IEEE Trans. Mobile Comput.*, vol. 7, no. 1, pp. 1–18, Jan. 2008.

[14] M. Guo, N. Pissinou, and S. S. Iyengar, "Pseudonym-based anonymity zone generation for mobile service with strong adversary model," in *Proc. Consum. Commun. Netw. Conf.*, 2015, pp. 335–340.

[15] F. Bonomi, R. Milito, P. Natarajan, and J. Zhu, "Fog computing: A platform for Internet Things analytics," *Big Data and Internet of Things: Roadmap for Smart Environments*. Springer, 2014, pp. pp. 169–186.

[16] F. Bonomi, "Connected vehicles, the Internet of Things, and fog computing," in *Proc. 8th ACM Int. Workshop Veh. Inter-Netw. (VANET)*, Las Vegas, NV, USA, 2011, pp. 13–15.

[17] L. M. Vaquero and L. Rodero-Merino, "Finding your way in the fog: Towards a comprehensive definition of fog computing," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 44, no. 5, pp. 27–32, Oct. 2014.

[18] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog computing and its role in the Internet of Things," in *Proc. 1st Ed. MCC Workshop Mobile Cloud Comput.*, 2012, pp. 13–16.

[19] S. J. Stolfo, M. B. Salem, and A. D. Keromytis, "Fog computing: Mitigating insider data theft attacks in the cloud," in *Proc. IEEE Symp. Secur. Privacy Workshops*, May 2012, pp. 125–128.

[20] M. Guo, X. Jin, N. Pissinou, S. Zanlongo, B. Carbunar, and S. Iyengar, "In-network trajectory privacy preservation," *ACM Comput. Surv. (CSUR)*, vol. 48, no. 2, p. 23, 2015.

[21] G. Han, W. Que, G. Jia, and L. Shu, "An efficient virtual machine consolidation scheme for multimedia cloud computing," *Sensors*, vol. 16, no. 2, p. 246, 2016.

[22] C.-Y. Chow, M. F. Mokbel, and X. Liu, "Spatial cloaking for anonymous location-based services in mobile peer-to-peer environments," *GeoInformatica*, vol. 15, no. 2, pp. 351–380, 2011.

[23] R.-H. Hwang, Y.-L. Hsueh, and H.-W. Chung, "A novel time-obfuscated algorithm for trajectory privacy protection," *IEEE Trans. Serv. Comput.*, vol. 7, no. 2, pp. 126–139, Jun. 2014.

[24] B. Palanisamy and L. Liu, "Attack-resilient mix-zones over road networks: Architecture and algorithms," *IEEE Trans. Mobile Comput.*, vol. 14, no. 3, pp. 495–508, Mar. 2015.

[25] C. 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, Jun. 2011.

[26] X. Wu and G. Sun, "A novel dummy-based mechanism to protect privacy on trajectories," in *Proc. IEEE Int. Conf. Data Mining Workshop*, Dec. 2014, pp. 1120–1125.

[27] C.-Y. Chow, M. F. Mokbel, and X. Liu, "A peer-to-peer spatial cloaking algorithm for anonymous location-based service," in *Proc. 14th Annu. ACM Int. Symp. Adv. Geograph. Inf. Syst.*, 2006, pp. 171–178.

[28] Y. Che, Q. Yang, and X. Hong, "A dual-active spatial cloaking algorithm for location privacy preserving in mobile peer-to-peer networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2012, pp. 2098–2102.

[29] S. Gao, J. Ma, C. Sun, and X. Li, "Balancing trajectory privacy and data utility using a personalized anonymization model," *J. Netw. Comput. Appl.*, vol. 38, pp. 125–134, Feb. 2014.

[30] T. Wang *et al.*, "Reliable wireless connections for fast-moving rail users based on a chained fog structure," *Inf. Sci.*, vol. 379, pp. 160–176, Feb. 2017.

[31] V. Cardellini, V. Grassi, F. L. Presti, and M. Nardelli, "On QoS-aware scheduling of data stream applications over fog computing infrastructures," in *Proc. Comput. Commun.*, 2015, pp. 271–276.

[32] M. Khari *et al.*, "Comprehensive study of Web application attacks and classification," in *Proc. 3rd Int. Conf. Comput. Sustain. Global Develop. (INDIACom)*, 2016, pp. 2159–2164.

[33] T. Hara, A. Suzuki, M. Iwata, Y. Arase, and X. Xie, "Dummy-based user location anonymization under real-world constraints," *IEEE Access*, vol. 4, pp. 673–687, 2016.

[34] J. C. Ho, C. H. Lee, and J. Ghosh, "Septic shock prediction for patients with missing data," *ACM Trans. Manage. Inf. Syst. (TMIS)*, vol. 5, no. 1, p. 1, 2014.

[35] L. A. Park, J. C. Bezdek, C. Leckie, K. Ramamohanarao, J. Bailey, and M. Palaniswami, "Visual assessment of clustering tendency for incomplete data," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 12, pp. 3409–3422, Dec. 2016.

[36] C. T. Tran, M. Zhang, and P. Andreae, "Multiple imputation for missing data using genetic programming," in *Proc. Annu. Conf. Genetic Evol. Comput.*, 2015, pp. 583–590.

TIAN WANG received the B.Sc. and M.Sc. degrees in computer science from Central South University in 2004 and 2007, respectively, and the Ph.D. degree from the City University of Hong Kong in 2011. He is currently an Associate Professor with National Huaqiao University, China. His research interests include wireless sensor networks, fog computing, and mobile computing.

JIANDIAN ZENG received the B.S. degree from Huaqiao University, China, in 2015, where he is currently pursuing the master's degree. His research interests include wireless sensor networks, mobile computing, and fog computing.

MD ZAKIRUL ALAM BHUIYAN (M'09) received the B.Sc. degree from the International Islamic University at Chittagong, Chittagong, Bangladesh, in 2005, and the Ph.D. and M.Eng. degrees from Central South University, China, in 2009 and 2013, respectively, all in computer science and technology. He was a Post-Doctoral Fellow with Central South University, a Research Assistant with The Hong Kong Polytechnic University, and a Software Engineer in industries. He is currently an Assistant Professor (research) with the Department of Computer and Information Sciences, Temple University. He is also a member of the Center for Networked Computing. His research focuses on dependable cyber physical systems, wireless sensor network applications, network security, and sensor-cloud computing. He is a member of the ACM. He has served as a Managing Guest Editor, the Program Chair, the Workshop Chair, the Publicity Chair, the TPC Member, and a Reviewer of international journals/conferences.

HUI TIAN received the B.Sc. and M.Sc. degrees from the Wuhan Institute of Technology, Wuhan, China, in 2004 and 2007, respectively, and the Ph.D. degree from the Huazhong University of Science and Technology, Wuhan. He is currently an Associate Professor with National Huaqiao University, China. His research interests include network and multimedia information security, digital forensics, and information hiding.

YIQIAO CAI received the B.S. degree from Hunan University, Changsha, China, in 2007, and the Ph.D. degree from Sun Yat-sen University, Guangzhou, China, in 2012. In 2012, he joined Huaqiao University, Xiamen, China, where he is currently a Lecturer with the College of Computer Science and Technology. He is interested in differential evolution, multi-objective optimization, and other evolutionary computation techniques.

YONGHONG CHEN received the B.Sc. degree from Hubei National University, and the M.Eng. and Ph.D. degrees from Chognqing University, Chongqing, China, in 2000 and 2005, respectively. He is currently a Professor of the College of Computer Science and Technology, Huaqiao University, Xiamen, China. His research interests include network security, watermarking, and nonlinear processing.

BINENG ZHONG received the B.Sc., M.Sc., and Ph.D. degrees in computer science from the Harbin Institute of Technology in 2004, 2006, and 2010, respectively. He is currently an Associate Professor with National Huaqiao University, China. His research interests include computer vision, pattern recognition, and mobile computing.