IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 17, 2019, accepted January 29, 2019, date of current version May 13, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2898996

An Attribute Generalization Mix-Zone Without Privacy Leakage

ZHANG LEI[1]*, HE LILI[1]*, LIU DESHENG[1], LI JING[1], JIANG QINGFENG[2], AND YUAN QI[3]

[1]College of Information and Electronic Technology, Jiamusi University, Jiamusi 154007, China
[2]College of Computer Science and Engineering, Changshu Institute of Technology, Suzhou 225500, China
[3]College of Communication and Electronic Engineering, Qiqihar University, Qiqihar 161006, China

Corresponding author: He Lili (8213662@163.com)

This work was supported in part by the University Nursing Program for Young Scholars with Creative Talents in Heilongjiang Province under Grant UNPYSCT-2017149, in part by the Special Doctor Scientific Research Fund Launch Project of Jiamusi University (Research on Privacy Protection of User Collaboration in Location Services), in part by the Natural Science Fund of Heilongjiang Province for Outstanding Youth (YQ2019F018), in part by the Basic Scientific Research Operating Expenses of Heilongjiang Provincial Universities and Colleges (Research on Wear Condition Monitoring technology of High Speed Milling Cutter Based on Deep Learning), in part by the Ministry of Education and Humanities Social Sciences Research Youth Fund Project under Grant 18YJCZH068, in part by the Natural Science Research Projects in Colleges and Universities of Jiangsu Province under Grant 18KJB520002, in part by the Research Start-Up Fund Project of the Changshu Institute of Technology under Grant KYZ2018005Q, in part by the National Natural Science Foundation of China under Grant 61872204, in part by the Research Project of Education Department of Heilongjiang Province under Grant 135309453, in part by the Foundation for Returnees of Heilongjiang Province of China under Grant LC2017027, and in part by the Jiamusi University Science and Technology Innovation Team Construction Project under Project CXTDPY-2016-3.

*Zhang Lei and He Lili are co-first authors.

ABSTRACT In the road network, mix-zone is usually considered as an efficient application that provides a balance between privacy protection and service quality. At the same time, the mix-zone can also resist the attack of tracking without any affection in the quality of feedback result of navigation. However, just as the old adage goes, nobody is perfect, the mix-zone also has two non-negligible deficiencies. The first one is that the mix-zone is difficult to cloak or generalize all types of the attributes emitted by the user. The other one is that the mix-zone is difficult to resist the attack initials by the disguiser who participates in the process of constructing the anonymous group. In order to cope with above-mentioned two problems, this paper provides an uncorrelated mix-zone based on the conception of attribute generalization and homomorphic encryption. In this algorithm, the process of attribute generalization is executed by an agency who is the winner in the private bidding, and the value of generalized attributes value is calculated by the secure multi-party computation under the principle of homomorphic encryption. Accordingly, the private information of the user in the mix-zone is preserved and any entities, no matter the agency or other participants, are prevented from any information about the user as there is no information been exposed. Furthermore, as the attribute of each user is generalized, for the mix-zone leavers, their attributes emitted are similar to each other; it will be difficult for the adversary to correlate any user with the attribute grasped before the user enter the mix-zone. Therefore, the potential hazard of adversary tracking the user with attribute comparison is solved. At last, in order to demonstrate the performance of our proposed algorithm in theoretical and practical fields, several analyses of security and simulation experiments are proposed, and the verification of security as well as the results of simulations and comparisons can further demonstrate the priority.

INDEX TERMS Road network, mix-zone, attribute generalization, homomorphic encryption.

I. INTRODUCTION

For the past few years, along with the development of position technology and wireless communication, location based service (LBS) has become prosperous and spread into the daily life of general public especially in vehicle network [1], [2]. This type of service can provide such as navigation, point of interesting request as well as commercial recommended etc to facilitate the routing or some other location-based activities, and then provide convenience for nearly all users. However, as more and more users became to utilize applications of this type of service, they had discovered a fact that the service may reveal the privacy of the user during the utilization. Because of that the user had to send her precise location along with the query for obtaining the feedback. More seriously, the

The associate editor coordinating the review of this manuscript and approving it for publication was Zhaolong Ning.

2169-3536 © 2019 IEEE. Translations and content mining are permitted for academic research only.
Personal use is also permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

exposure of precise location can even incur potential physical threats such as tracking, robbery or some other crimes. For fearing of exposing the personal privacy, some persons became to reduce the ratio of utilizing LBS, and someone even attempted to avoid utilization, which will restrict the development and promotion of LBS.

In order to cope with the problem of privacy leakage, a great many of researchers had provided their conceptions. Among them, the most deeply influenced conception is the k-anonymity which proposed by Gruteser and Grunwald [3]. In this conception, the real location of the user has to be mixed with other $k-1$ locations and sends to the LBS provider simultaneously, so the real location will be generalized with other k-1 locations and the probability of the adversary successfully identify the real one will be reduced, the uncertainty of the adversary will be promoted, so as to protect the personal privacy of the user. Then following the conception of k-anonymity, such as conceptions of region generalization [4], location diversity [5] and queries diversity [6] etc are provided successively, and gradually became the conception of snapshot queries protection [7]–[9] and continuous queries protection [10], [11] in Euclidean space. But it is difficult to simply utilize these methods to the real environment, as the real environment of the user is composed by various types of roads, and these roads are intertwined with each other. If simply utilize these methods used in Euclidean space, the anonymous location may be located in a position that cannot be reached in any road or located in the position which can be identified by the adversary easily, then the anonymous location will affect these methods and invalid to protect the personal privacy. Therefore, the researcher has to provide methods that can be used in the road network to adapt the real environment [12]–[14]. However, such methods are failure in the ability of resisting the attack of tracking as well as in the exorbitant affection on the quality of service, which brings about another type of protection scheme called mix-zone becomes prosperously and been deeply researched [15]–[20]. In general, mix-zone can be seen as an area of black box, in this region, users can communicate and exchange information with each other, but no information can be detected out of this region, so the user can exchange the pseudonym and other identifications with each other to obfuscate the tracking and generalize the identifications that the adversary can be used. Furthermore, as the mix-zone is deployed in some critical areas that to obfuscate the tracking, it is not deployed in all of road network, which means the protection algorithm does not be executed all the time and will just lead weaker affection to the service of LBS. Thus, the series of advantages makes the mix-zone will be better to be used in the road network than other methods.

However, as attributes of the user not just as the pseudonym, the query interval and query content, some unconcerned attributes can also be utilized to correlate the real user, so just conceal or generated these limited number of attributes is not enough and above methods of mix-zone will be invalid when the adversary utilize these attributes. As the viewpoint of Lei *et al.* [21], the adversary can utilize nearly all types of attributes and correlates the real user with attributes correlation, and then guesses the location of the user to obtain the personal privacy. Although several algorithms used to generalize attribute had been proposed in the past few years [22]–[24], they are not satisfied to be simply used in mix-zone, as mix-zone is vulnerable to the adversary involved in the anonymous group and there is no trusted user in mix-zone. Furthermore, mix-zone is also provided not only to reduce the attribute correlation but also facilitate the affection of quality of service, the methods acting along the road is also not suitable in service quality let alone the vehicle network [25]–[27]. Thus, in order to cope with attribute correlation and cope with the un-trusted user in mix-zone, a mix-zone called homomorphic encryption based attribute generalization (short for HEBAG mix-zone) is proposed. This mix-zone is based on the conception of attribute generalization, and then with the help of homomorphic encryption to bid the agency to calculate the similar attributes, so there is no information revealed to any type of entities. Accordingly, the disguiser cannot obtain any privacy of the user no matter that he is participated in the mix-zone. When users leave the mix-zone, each user utilizes the transformed attributes calculated in the mix-zone, so similarity of attributes with each other, the adversary is difficult to identify any user with attribute correlation, and then the personal privacy is protected. The main contributions can be summarized as follows.

1) We proposed an attributes generalization mix-zone with homomorphic encryption, and this type of mix-zones can resist the attack of tracing with user's attributes and prevent privacy leakage during the process of information transformation by collaborative users.
2) We proposed a private competitive bidding comparison algorithm, and with this algorithm an agency was elect to dispose the attributes generalization, but no additional information can be leaked.
3) We proposed an attributes generalization algorithm with homomorphic encryption, with this algorithm there is not any private information can be detected by other users including the agency.
4) We evaluated the proposed algorithm in experimental verification, so the ability of privacy protection as well as the efficiency of algorithm execution is compared with other similar algorithms.

The rest of this paper is organized as follows. Firstly, we introduced preliminaries in Section 2. Section 3 presents the scheme of homomorphic encryption based attribute generalization mix-zone. Then we show the results of experimental with cause analysis in Sections 4. Finally, we draw conclusions and future works in Section 5.

II. PRELIMINARY

A. PRIVACY THREAT

In general, mix-zone is mainly used to resist the attack of tracking and utilizes the uncertainty in mix-zone to cut off the

correlation between the users of entering and leaving. But as the variety of attributes emitted from each user, an adversary can utilize the difference to identify the real user and continuous tracking the user along the road. So just generalize a limited number of attributes (such as pseudonym and query interval) is not enough. Suppose that the amount number of attributes of each user can be defined as trilateral, pentagon and star and shown in figure 1. The adversary can utilize the different attributes of each user to identify the tracked user, as the mix-zone cannot conceal all attributes and it is failed to protect the personal privacy. Just as shown in figure 1, each user still maintains the shape of its attributes and there is no change in the set of attributes.

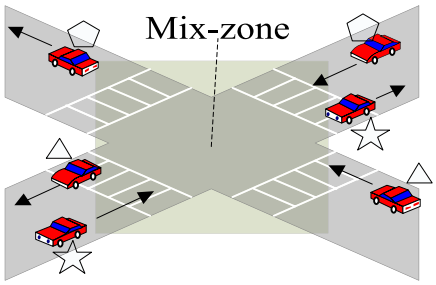


FIGURE 1. The sum attributes of each user outside the mix-zone.

Based on the analysis of figure 1, the attack of correlation initiated by the adversary can be denoted as the comparison of attributes similarity. Suppose that the set of attributes about a tracked user is $A_s = (a_{s,1}, a_{s,2}, \ldots, a_{s,m})$, where m is the number of attributes that the adversary can utilize to correlate this user. The set of attributes that the user leaves the mix-zone can be denoted as $A'_s = (a'_{s,1}, a'_{s,2}, \ldots, a'_{s,m})$, if A_s and A'_s satisfy $sim(A_s, A'_s) < \lambda$, then the adversary can regard that these attributes are emitted by the same user and users of entering and leaving can be merged into the same user, he can continue tracking this user to get the privacy. Where m is the number of attributes that the adversary can utilize, λ is the threshold value to estimate whether the attribute set similar with the tracked user.

B. CONCEPTION AND BASIC IDEAS

In order to cope with the attack of similarity attribute comparison, the most efficient scheme is to generalize the set of attributes, and utilize the similar attributes to obfuscate the correlation. The conception of attribute generalization is to share the similar attribute set that nearest to the original one, and utilize the distribution of similar attributes to promote the uncertainty of the adversary. In figure 2, the emitted set of attributes about each user is transformed and changed into similar attributes set. As all users are emitting pentagons after leaving the mix-zone, the adversary cannot identify the user with attribute comparison, and then the personal privacy of the user is protected.

From figure 2, we can see that the emitted set of attributes about each user is transformed and changed into pentagons after leaving the mix-zone, which can be formalized as

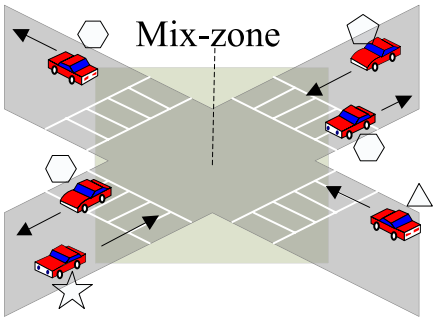


FIGURE 2. The transformed and changed attributes of each user.

$sim(A_s, A'_n) < \lambda, (s \in n)$, where n is the number of users in current mix-zone. However, who disposes the transformation of attributes and whether the other user can get the attributes of the user still unsolved, as the disguiser can be involved in the mix-zone to contact with the process of attribute generalization. Thus, the conception of attribute generalization also has to consider the security of generalization procedure. Then based on the conception of security generalization and utilize the characteristic of homomorphic encryption, this paper proposes two types of secure computation algorithms to achieve private agency bidding and secret calculating for similar attributes. During the procedure of attribute generalization, there is no information that can be identified by any entities except the user itself. Furthermore, the generalized attributes after leaving the mix-zone can also obfuscate the identification of the adversary.

III. THE SCHEME OF HOMOMORPHIC ENCRYPTION BASED ATTRIBUTE GENERALIZATION MIX-ZONE

The procedure of our attribute generalization mix-zone depends on the confidence of private competitive bidding as well as private calculation of similar attributes. So the procedure can be divided into two phases called private competitive bidding comparison as well as private calculation of similar attributes. In the following, these two phases will be elaborated respectively.

A. THE PHASE OF PRIVATE COMPETITIVE BIDDING COMPARISON

As the capability of movable device is restricted, users have to pay feedback to the user who disposes the process of attribute calculation, so the agency will be the user who offers the lowest price in this zone. In order to guarantee the fairness, the bidding process in this mix-zone needs to be secret to others. In addition, the bidding winner called agency has to satisfy two basic conditions: the first one is the agency has the ability to communicate with users in current mix-zone showed in figure 3; the other one is the agency has the lowest bidding value in this zone. For calculating the lowest bidding, every two users had to compare the bidding with each other and the process of comparison can be shown in figure 4.

Based on the above analysis, the bidding comparison can be converted into the problem of private comparison.

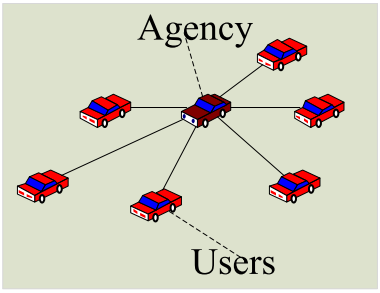


FIGURE 3. The communicated ability of agency in mix-zone.

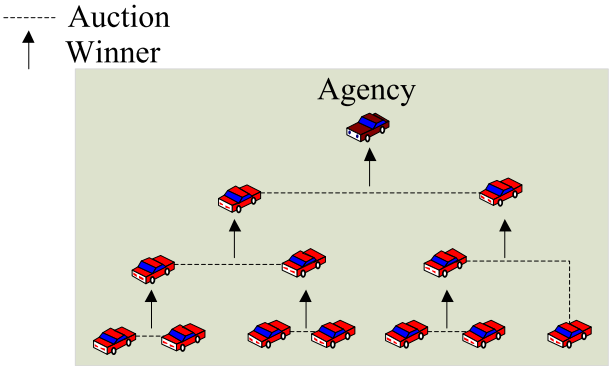


FIGURE 4. The procedure of choosing agency with bidding comparison.

Then with the conception of literature [28], we will have the following procedure of private comparison.

Suppose there are n users in the current mix-zone, and each user $P_i(1 \leq i \leq n)$ has the bidding value which length is l bit. The bidding values can be denoted as $b_1, b_2, \ldots, b_n$, when converted to binary, it can be denoted as $\vec{b_i} = (b_{i,l-1}, b_{i,l-2}, \ldots, b_{i,0})$. Thus, the bidding value $b_1, b_2, \ldots, b_n$ can be compared with the following steps.

Step 1: $P_i(1 \leq i \leq n)$ generates the key pair of public and private (pk_i, sk_i) based on fully homomorphic encryption, and sends the public key pk_i to each user located in current mix-zone respectively.

Step 2: P_1 utilizes the public key pk_1 to encrypt the binary bidding $\vec{b_1} = (b_{1,l-1}, b_{1,l-2}, \ldots, b_{1,0})$ bit by bit, and gets the encrypted information $E(b_1) = (E(b_{1,l-1}), E(b_{1,l-2}), \ldots, E(b_{1,0}))$, and then sends the bidding information $E(b_1)$ to other $n-1$ users $P_k(2 \leq k \leq n)$.

Step 3: If P_k receives the encrypted bidding $E(b_1)$, she first utilizes pk_1 to encrypt the bidding of herself $\vec{b_k} = (b_{k,l-1}, b_{k,l-2}, \ldots, b_{k,0})$, and gets the encrypted information $E(b_k) = (E(b_{k,l-1}), E(b_{k,l-2}), \ldots, E(b_{k,0}))$, then calculates

$$\begin{aligned}\bar{e} \triangleq E(c_k) &= (E(c_{k,l-1}), E(c_{k,l-2}), \ldots, E(c_{k,0})) \\ &= (E(c_{k,l-1}) \odot E(1), E(c_{k,l-2}) \\ &\quad \odot E(1), \ldots, E(c_{k,0}) \odot E(1)).\end{aligned}$$

where $E(1)$ is the ciphertext of 1 that encrypted by fully homomorphic encryption, and $E(c_k) \odot E(1)$ denoted the result of $E(c_k) + E(1) \bmod d$, which means $E(c_k) \odot E(1)$ is the bidding shifting l bits added with the modular of $E(1)$ in d bits. If the result exceeds the range of l bits only needs to remove the value lower than l bits.

After finishing above calculation, P_k calculates $ans(E(b_1), E(b_k)) = E(b_1) + \bar{e}E(g_k)$. Where $E(b_1) + \bar{e}$ denotes the addition of $E(b_1)$ and $E(c_k)$ within l bits. $E(g_0) = E(d_{1,0}) \odot E(c_{k,0})$, $E(t_0) = E(d_{1,0}) \cdot E(c_{k,0})$, $E(g_i) = E(d_{1,i}) \odot E(c_{k,i}) \odot E(t_{i-1})$, $E(t_i) = (E(d_{1,i}) \cdot E(c_{k,0})) \odot (E(d_{1,i}) \cdot E(t_{i-1})) \odot (E(c_{k,0}) \cdot E(t_{i-1}))$, $1 \leq i \leq l-1$. $E(g_i), E(t_i)(0 \leq i \leq l-1)$ are results of the $i$bit carry to the $i+1$ bit, $E(d_{1,i}) \cdot E(c_{k,0})$ denotes the modular of $E(d_{1,i})$ and $E(c_{k,0})$ in d bits multiplication. In addition, the addition and multiplication of d bits are all according to the procedure of fully homomorphic encryption, and between the procedure of addition and multiplication the algorithm executes homomorphic decryption to reduce the noise generated in above process. Then P_k gets $ans(E(b_1), E(b_k))(E(g_{k,l-1}), \ldots, E(g_{k,0}))$ and sends to $E(g_{k,l-1})$ user P_1.

Step 4: P_1 decrypts $E(g_{k,l-1}), 2 \leq k \leq n$ and gets $g_{k,l-1}$, if $g_{k,l-1} = 0$, it means $b_1 \geq b_k$, otherwise $g_{k,l-1} = 1$ and $b_1 < b_k$, then P_1 sends the result to P_k.

In the similar way, after $n(n-1)/2$ rounds of security comparison, the minimum value of bidding will be produced. Then the user who puts forward this value can be considered as the agency to calculate the attribute generalization.

B. THE PHASE OF PRIVATE CALCULATION OF SIMILAR ATTRIBUTES

In general, the attribute set of a moving user in the road can be denoted as $A = (a_1, a_2, \ldots, a_m)$. For at least n users in the mix-zone, this set can be denoted as $A_i = (a_{i,1}, a_{i,2}, \ldots, a_{i,m}), 1 \leq i \leq n$. The aim of attribute generalization is to generalize the attribute set of each user similar with each other, so as each attribute set satisfies $A_1 \cong A_2, \ldots, \cong A_n$. Suppose for arbitrary two chosen users u_i and u_j, the attribute set satisfies $A_i - A_j \leq \lambda$, where λ is the threshold of the adversary difficult to identify. In order to achieve this type of attribute similarity as well as to preserve privacy for each user from being obtained by the agency and other disguisers, a similar attributes calculation method has been proposed based on the conception of preserving meeting location determination [29]. This method can complete the calculation of similar attributes just in one round. The procedure of this method deals with message can be described as follows.

Before sending any attribute to the agency, the agency first utilizes ElGamal to generate her public key $pk_s = \{p, g, Z\}$, and the private key is $sk_s = \alpha$. Where p is a large prime number such that $\|p\| = 1024$, g is the generator of $\mathbb{Z}_p$b and $Z = g^{\alpha} \bmod p$, $\|\cdot\|$ indicates the bit length. Then the agency publishes the public key $pk_s = \{p, g, Z\}$ to all users in the mix-zone.

Step 1: For two arbitrary chosen users u_i and u_j, they satisfy $i < j$. The agency securely distributes the secret random numbers $r_{ij,1} \in \mathbb{Z}_p^*$, $r_{ij,2} \in \mathbb{Z}_p^*$, $r_{ij,3} \in \mathbb{Z}_p^*$ and $r_{ij,4} \in \mathbb{Z}_p^*$ to user u_j, and at the same time distributes the secret random

numbers $r'_{ij,1} \in \mathbb{Z}_p^*$, $r'_{ij,2} \in \mathbb{Z}_p^*$, $r'_{ij,3} \in \mathbb{Z}_p^*$ and $r'_{ij,4} \in \mathbb{Z}_p^*$ to user u_i, where $r_{ij,t} + r'_{ij,t} = r_{ij}(t = 1, 2, 3, 4)$. Then the user u_i randomly chooses a positive number $s_i \in \mathbb{Z}_p^*$ that satisfies $\|s_i\| \leq \|p\| - 1$ and keeps s_i secret. All users in current mix-zone secretly share a common random positive number $r \in \mathbb{Z}_p^*$, where r satisfies $\|r\| \leq \|p\| - 35$.

Step 2: For the user u_i, she divides the set of attributes into two arbitrary portions and transmits the attribute set into $A_i = (x_i, y_i) \in (\mathbb{Z}_p^*)^2$. For the user u_j, the set is transmitted into $A_j = (x_j, y_j) \in (\mathbb{Z}_p^*)^2$ and u_j calculates the following ciphertexts as follows.

$$\begin{aligned} C_{j,1} &= r(x_j^2 + y_j^2)Z^{r_{ij,1}} \bmod p \\ C_{j,2} &= Z^{r_{ij,2}} \bmod p \\ C_{j,3} &= 2rx_jZ^{r_{ij,3}} \bmod p \\ C_{j,4} &= 2ry_jZ^{r_{ij,3}} \bmod p \\ C_{j,5} &= g^{r_{ij,4}} \bmod p \end{aligned}$$

Then u_j sends the ciphertexts $(C_{j,1}, C_{j,2}, C_{j,3}, C_{j,4}, C_{j,5})$ to the user u_i, and then u_i generates the following ciphertexts upon receiving the ciphertexts $(C_{j,1}, C_{j,2}, C_{j,3}, C_{j,4}, C_{j,5})$.

$$\begin{aligned} C_{ij,1} &= C_{j,1}Z^{r'_{ij,1}} + C_{j,2}r(x_i^2 + y_i^2)Z^{r'_{ij,2}} \bmod p \\ &= r(x_i^2 + y_i^2 + x_j^2 + y_j^2)Z^{r_{ij}} \bmod p \\ C_{ij,2} &= (C_{j,3}x_i + C_{j,4}y_i)Z^{\tilde{r}_{ij,3}} \bmod p \\ &= 2r(x_ix_j + y_iy_j)Z^{r_{ij}} \bmod p \\ C_{ij,3} &= C_{j,2}s_iZ^{\tilde{r}_{ij,2}} \bmod p = s_iZ^{r_{ij}} \bmod p \\ U_{ij} &= C_{ij,1} - C_{ij,2} + C_{ij,3} \\ &= [r((x_i - x_j)^2 + (y_i - y_j)^2) + s_i]Z^{r_{ij}} \bmod p \\ V_{ij} &= C_{j,5}g^{r'_{ij,4}} = g^{r_{ij}} \bmod p \end{aligned}$$

At last, the user u_i sends the ciphertexts (U_{ij}, V_{ij}) to the agency.

Step 3: Upon receiving the ciphertexts (U_{ij}, V_{ij}) from the user u_i, the agency decrypts it with the private key $sk_s = \alpha$ and obtains the plaintext

$$\begin{aligned} m_{ij} &= r((x_i - x_j)^2 + (y_i - y_j)^2) + s_i \bmod p \\ &= rd_{ij}^2 + s_i \bmod p, \end{aligned}$$

where $d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ denotes the attributes distance between two users u_i and u_j. As $\|r\| \leq \|p\| - 35$, $\|s_i\| \leq \|p\| - 1$ and $0 < rd_{ij}^2 + s_i < p$, we can conclude $0 \leq m_{ij} = rd_{ij}^2 + s_i < p$.

Step 4: For each user in the mix-zone, the agency calculates the mean value of m_{ij} by

$$\bar{m}_i = \frac{\sum_{j=1,j\neq i}^{n} m_{ij}}{n-1} = r\frac{\sum_{j=1,j\neq i}^{n} d_{ij}^2}{n-1} + s_i = r\bar{d}_i^2 + s_i,$$

and then the agency calculates the difference between them with $r\Delta_{ij} = |m_{ij} - \bar{m}_i| = |rd_{ij}^2 - r\bar{d}_i^2| = r|d_{ij}^2 - \bar{d}_i^2|$. From each difference $r\Delta_{ij}, (j \in [1, n], j \neq i)$, the agency just has to calculate r times of the difference of each pair of attributes to obtain the attribute distance variance σ_i by $r\sigma_i = \frac{\sum_{j=1,j\neq i}^{n} r\Delta_{ij}}{n-1}$. After obtaining the values $r\sigma_i (i \in [1, n])$, the agency chooses the minimum value of $r\sigma_k (k \in [1, n])$ as well as the index of corresponding k. As r is a positive value and σ_k is the minimum value among $\{\sigma_1, \sigma_2, \ldots, \sigma_n\}$, which means the result of σ_k is the value that similar to each user and the attribute A_k is the attribute set that satisfies $A_1 \cong A_2, \ldots, \cong A_n$ and each user can utilize this value to instead of the attribute value when leaving the mix-zone. In addition, during the whole procedure of similar attributes calculating, all the information transformed is encrypted, no matter the users or the agency. Each entity just knows the value of generalized result, as the sends message to the agency is encrypted. Consequently, the procedure of this algorithm can be summarized as follows, firstly the users in mix-zone had bidding for the agency; secondly the agency disposed the attribute information for each other; thirdly each user utilized the attribute set instead of the original attributes, and then each user will show the similar attribute so that the adversary will be difficult to identify the real user when all users leave the mix-zone.

C. SECURITY ANALYSIS

As the conception of our proposed mix-zone is attribute generalization, and the conception to resist disguiser is based on homomorphic encryption, so this mix-zone can provide privacy protection from two types of adversary called active adversary and passive adversary. The active adversary can track the user along the moving with the similarity of attributes, except the user located in the mix-zone. The passive adversary can be involved in the mix-zone and obtains the information of the user through the process of attribute generalization. Thus, the security of our proposed mix-zone has to consider two aspects. The first one is the uncertainty of the adversary guesses the real user by comparing the similarity of attributes, the other one is whether the procedure of attribute generalization can leak the information of the user to the agency or other users in the mix-zone. In addition, each aspect has to utilize a specific metric to measure the privacy or information leakage.

For measuring the uncertainty of the adversary guesses the real user with attribute similarity, entropy is used as an efficient metric. Suppose the probability of an adversary successfully identifies the real user is $p_i, (1 \leq i \leq n)$, where n is the number of users located in mix-zone. Then the uncertainty of this adversary utilizes attribute comparison to identify the real user when leaving the mix-zone can be denoted by entropy and described as $H_i = -\sum_{i=1}^{n} p_i \log_2 p_i$. According to the theorem of Jaynes maximum value of entropy, the larger value of H_i means the higher uncertainty of the adversary and the more difficult of the adversary identifies the real user from the group of users in mix-zone. In our proposed mix-zone, suppose the set of attributes of a tracked user can be denoted as $A = (a_1, a_2, \ldots, a_m)$, u_i and u_j are two arbitrary

chosen users from the group of at least n users in mix-zone, and attributes of these two users are $A_i = (a_{i,1}, a_{i,2}, \ldots, a_{i,m})$ and $A_j = (a_{j,1}, a_{j,2}, \ldots, a_{j,m})$ respectively. As the adversary utilizes attributes similarity to identify the real user, so the probability calculated from the tracked user to these two users can be denoted as $sim(A, A_i)$ and $sim(A, A_j)$, and then the adversary can utilize the value which higher than λ to guess the user as the tracked user. Suppose $p_{u,i}$ and $p_{u,j}$ are values of the probability of an adversary guesses the real user with two arbitrary chosen users u_i and u_j and $H = -p_{u,i} \log_2 p_{u,j} - p_{u,j} \log_2 p_{u,i}$ is the entropy used to measure the uncertainty of an adversary. Without attribute generalization or just few attributes are generalized, the value H will be less than 1 or even equivalent to 0. But with attribute generalization for nearly all types of attributes, the value of similarity of current attributes to the tracked user will satisfy $sim(A, A_i) \cong sim(A, A_j) < \lambda$, where λ is the threshold value used by an adversary successfully guesses the real user. In addition, as the set of attributes are similar to each other, the probability of these two arbitrary chosen users will satisfy $p_{u,i} = p_{u,j}$, according to the theorem of Jaynes, the value of H will be higher than any others, which means the highest uncertainty for an adversary successfully identify the real user from the group of users located in the mix-zone and he is difficult to track the real user.

As two phases of our proposed mix-zone has utilized encryption algorithm to deal with private competitive bidding as well as private attributes calculation, so the security of encryption algorithm used in these two phases has to be discussed. Firstly, for the private competitive bidding, each user located in the mix-zone has to bid for becoming agency to dispose the attribute generalization. If and only if the bidding value is the minimum, the user can be selected as the agency and gets the compensation. During the procedure of bidding, all bidding values are encrypted nobody can learn any bidding information about any other. However, as the agency is selected from n users located in the mix-zone, some users may collude with others to infer the minimum bidding. But in our proposed bidding algorithm, all users just learn the ranking of him but not the exact value of bidding, so the adversary needs at least $n-1$ users $P_i (1 \leq i \leq n-1)$ collude with others to infer the real bidding. But in mix-zone, users of the group are selected arbitrarily and it is difficult to be involved in at least $n-1$ users can collude with each other. For another, the agency is just paid for a limited number of compensation, and the cost for at least $n-1$ users is higher than the compensation, which means the compensation maybe much less than the cost of collusion. Thus, we can conclude that the procedure of private competitive bidding is secure and no bidding information can be revealed to any entities.

For the phase of private calculation of similar attributes, the security has to consider two aspects. The first one is the secret of attribute information to agency, the other one is the secret of attribute information to users in mix-zone. For the agency, as the agency has to decrypt the ciphertexts (U_{ij}, V_{ij}) with the private key $sk_s = \alpha$ and obtains the plaintext m_{ij}. But as the plaintext

$$\begin{aligned} m_{ij} &= r((x_i - x_j)^2 + (y_i - y_j)^2) + s_i \bmod p \\ &= rd_{ij}^2 + s_i \bmod p \end{aligned}$$

is calculated by modular arithmetic, which means the value before the modular arithmetic will be multiple and difficult to confirm the exact value. Furthermore, as the plaintext is constituted with the attributes of users u_i and u_j, the agency just obtains the sum of two portions of attribute value but not the exact value, so the agency is unable to get any attribute information about any users, no matter he is colluded with other users.

For other users except the agency, suppose ε is the private parameter, A is the adversary, B is the user who is tracked before entering the mix-zone, and the attacks for obtaining attributes information can be transmitted into the Game_A^B between A and B. In the procedure of attribute generalization, suppose A can obtain the private key of the agency sk_s and all the public parameters of B. All sets of attributes of users located in the mix-zone can be denoted as $A_1, A_2, \ldots, A_n$. B excuses the algorithm of private attribute generalization and gets the generalized attributes $f(A_{gen}) = g(f(A_1), \ldots, f(A_n))$, and then she sends the ciphertext $f(A_1), \ldots, f(A_n)$ as well as the index related to A_{gen} to A, and then B randomly chooses attributes $A_r = (x_r, y_r)$ and select arbitrary number $\varepsilon \in \{0, 1\}$, if $\varepsilon = 0$, user B A_{gen} to A, otherwise sends A_r to A. The adversary A guesses the value of private parameter $\varepsilon' \in \{0, 1\}$, if and only if $\varepsilon' = \varepsilon$, the adversary A wins the game Game_A^B. Then the advantage of the adversary A can be denoted as $Adv_A = 2pr[\varepsilon' = \varepsilon] - 1$. As the adversary A cannot utilize the similarity comparison to successfully guess the exact value of ε', which means the probability of A guesses the value of private parameter satisfies $\varepsilon' = \varepsilon$. Thus, the advantage of the adversary A can be calculated as $Adv_A = 2pr[\varepsilon' = \varepsilon] - 1 = 2 \times \frac{1}{2} - 1 = 0$, which means no advantage. Accordingly, we can conclude that no attribute information can be obtained by other users.

For time complexity, although agency is added in the procedure of information imposing, it does not bring any influence. This is because of that the whole procedure is affected by the number of attributes as well as the anonymous value that the user chosen, so the time complexity can be considered as the result of $n \times k$, thus we can conclude that the time complexity of this algorithm is $O(nk)$, where n denotes the number of attributes and k denotes the anonymous value.

In conclusion, based on above analyses, we can conclude that our proposed mix-zone can provide attribute generalization without leaking any attribute information for any entities, so this mix-zone can resist attacks from both of the active and passive adversaries and can cope with attacks of attributes correlation as well as disguisers monitor in the same mix-zone.

IV. EXPERIMENTAL VERIFICATION

Another way to verify the performance of our proposed algorithm is deployed it in experimental environment and

compared with other similar algorithms. In this paper, the central part of BerlinMOD Data is used and we choose several intersections as the location to deploy the mix-zones. The simulation experiment is implemented on a laptop with Intel Core i7 1.70 GHz CPU, 8 GB RAM memory and Windows 10×64 ultimate operating system, and then Matlab R2017a is used as the instrument to verify the performance. Furthermore, in order to further express the superiority of our proposed mix-zone, several similar algorithms are used to compare with our proposed scheme. These algorithms include the mix-zone that utilizes time delay to generalize the query interval (delay-tolerant mix-zone) [30], the mix-zone that utilizes shape distortion to reduce the correlation (shifted mix-zone) [31], the mix-zone that utilizes multiple dimensions to obfuscate the correlation (multiple mix-zone) [19] as well as the mix-zone that utilizes identity-based authentication to conceal the attributes (cryptographic mix-zone) [20]. The comparison will focus on the ability of privacy protection and the efficiency of algorithm execution. For measuring the ability of privacy protection, entropy is used to verify the level of attribute generalization, the effect of attribute similarity utilizes attribute differential to measure, at last the pair entropy is used to measure the correlation degree between two arbitrary chosen users entering and leaving the mix-zone. For measuring the efficiency of each algorithm, the running time of accomplishing mix-zone and achieves attribute generalization as well as the success ratios of executing an algorithm are used. For alleviating the affection of parameters, we set the maximum number of users and the number of attributes as 30, if the number of users is changing we set the number of attributes as 15, if the number of attributes changing we set the number of users as 10.

From figure 5, we can see that entropy values of all algorithms are descending along with the increasing of attributes except our proposed mix-zone. The reason for this phenomenon can be ascribed as the HEBAG mix-zone is designed for generalizing nearly all types of attributes, the number of attributes that can be dealt is much higher than any other algorithms, so the privacy protection level does not be affected by the increasing number of attributes. For other algorithms, as some attributes are encrypted, the performance

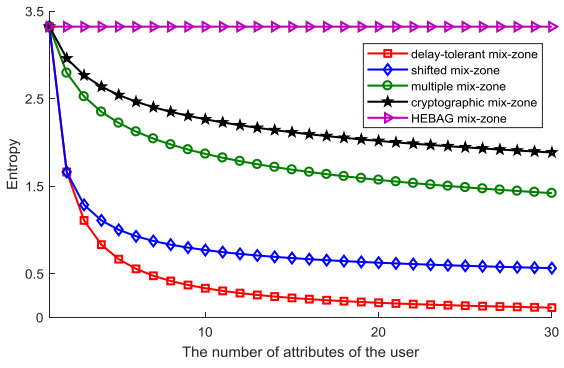


FIGURE 5. The value of entropy changed with attributes number (the number of users is 10).

of cryptographic mix-zone is better than the rest of other algorithms, but the number of attributes that can be concealed is limited, so the entropy value is lower than HEBAG mix-zone. As multiple mix-zones utilize more numbers of mix-zones to deal with correlation, which makes the performance better than the shifted mix-zone and the delay-tolerant mix-zone, but the ability is lower than encryption. At last, algorithms of shifted mix-zone and the delay-tolerant mix-zone are designed to generalize the attribute of moving speed and query interval, the limitation of attribute number descends the value of entropy, but as the shifted mix-zone can involve more regions than the delay-tolerant mix-zone, its performance is better than delay-tolerant mix-zone.

From figure 6, in the condition of the number of attributes is constant, the value of entropy is ascending with the increasing of users' number, which means the uncertainty of the adversary is continuous ascending. This is because of that the probability of the adversary guesses the real use depends on the number of users, the more users the more difficult for guessing. Among these algorithms, the HEBAG mix-zone can acquire the maximum value of entropy, as this mix-zone can generalize nearly all types of attributes, and nearly no attributes can be used to correlate any user. For other algorithms, the algorithm of multiple mix-zones utilize more numbers of mix-zones, which can generalize more attributes with different mix-zones, so the entropy value is just a bit lower than HEBAG mix-zone. Although the cryptographic mix-zone utilizes the encryption to generalize the attribute, the number is limited sometimes even less than 15, so its performance is worse than multiple mix-zones. At last, as shifted mix-zone can cover more regions than delay-tolerant mix-zone, its performance is better than delay-tolerant mix-zone, but still affected by the limitation of shifting distance.

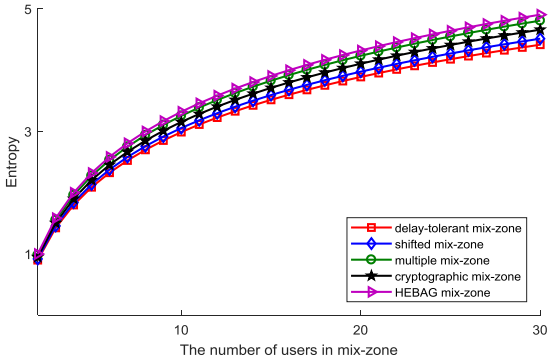


FIGURE 6. The value of entropy changed with users number (the number of attributes is 15).

From figure 7, we can see the differential of attributes is ascending with the increasing of attribute number, except the HEBAG mix-zone. In addition, the larger number of attributes means the more differential in attributes. The reason for this phenomenon can be ascribed as the number of attributes that can be disposed by algorithms, the more number of attributes can be disposed the less of this value.

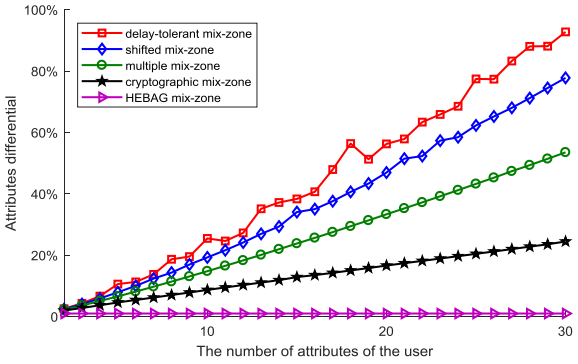


FIGURE 7. The differential of attributes changed with the number of attributes.

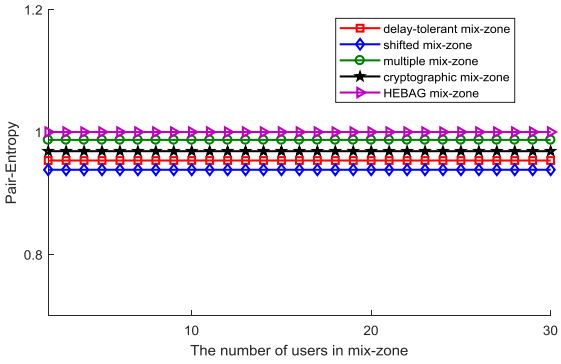


FIGURE 9. The pair-entropy changed with the number of users (the number of attributes is 15).

So in this figure, along with the number of attributes that being disposed is limited, the differential of other algorithm is expanding, but not affects the HEBAG mix-zone, as it can dispose nearly all types of attributes.

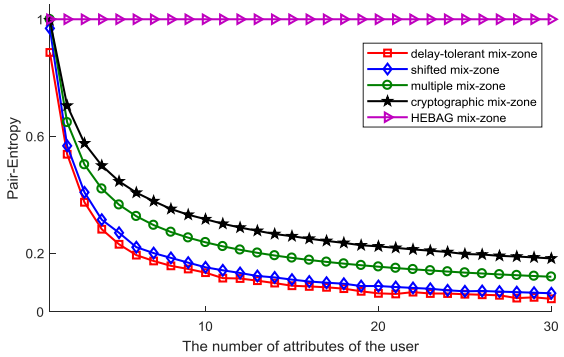


FIGURE 8. The pair-entropy changed with the number of attributes (the number of users is 10).

For measuring the correlation between the entering and leaving users, the pair-entropy is utilized, and the higher of the value of pair-entropy the more difficult for the adversary to correlate these two users. In figure 8, we can see that the HEBAG mix-zone can reach the maximum value of pair-entropy, as this algorithm can conceal nearly all types of attributes, so the user who leaves the mix-zone is similar to others and difficult to be correlated with the user who entered the mix-zone. However, the other algorithms are all restricted by the number of attributes that can dispose, and some attributes that can be used to correlate the user, which means the user has a higher probability to be identified by the adversary.

From figure 9, we can see that along with the increasing of users' number, the value of pair-entropy is unchanged and presents a linear tendency. This is because of the entering and leaving users are determined in our experiment, but the pair-entropy is mainly used to measure the degree of correlating the pair of entering and leaving users with attribute similarity and this tendency does not affect by the increasing number of other users. However, among these leaner values, the maximum value is generated by HEBAG mix-zone, as this algorithm nearly generalizes all types of attributes and the probability of correlating these two users is difficult, even if the pair of users is determined. For other algorithms, as the disposed number of attributes is restricted, the value of pair-entropy is lower than HEBAG mix-zone, and the less number of attributes the lower value of pair-entropy.

From figure 10, we can see that the running times of all algorithms are ascending with the increasing of attributes except HEBAG mix-zone, as this mix-zone can dispose generalization to nearly all attributes in various types in just one round calculation rather than dispose attributes with the increasing number, so it's running time does not be changed with the increasing of attribute number. For other algorithms, as the procedure of attributes disposition depends on the performance of dealing with the category of attributes that appeared by the user, the procedure of disposition has to deal with each attribute, so the running time is ascending with the increasing number of attributes. Among these algorithms, running times of cryptographic mix-zone and multiple mix-zones are higher than others, as these two algorithms had to encrypt each attribute or utilize multiple regions to cope with each attribute, which aggravate the complexity of attribute disposition. At last, running times of delay-tolerant mix-zone and shifted mix-zone do not sharply ascending with

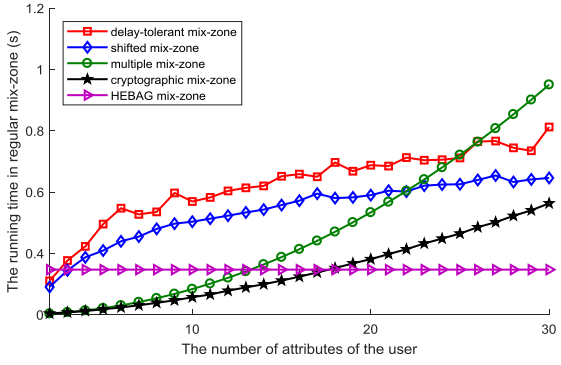


FIGURE 10. The running time in regular mix-zone changed with the number of attributes (the number of users is 10).

the increasing of attribute number like cryptographic mix-zone and multiple mix-zones, this is because of that these two algorithms do not need to repetitive execute the algorithm to dispose each increasing attributes. Furthermore, the number of attributes that these two algorithms can deal with is limited, and sometimes the algorithm does not run enough time to deal with the procedure of attribute generalization, which leads the limited increasing of running time.

From figure 11, we can see that different from the running time shown in regular mix-zone, the running time of HEBAG mix-zone in irregular mix-zone is much higher than the running time in regular mix-zone. The reason for this phenomenon can be ascribed as the procedure of selecting users, as the irregular mix-zone may involve more users than regular mix-zone and more users bidding for the agency will occupy more running times. For other algorithms, the performance of shifted mix-zone in running time in irregular mix-zone is better than the remaining algorithms, as this mix-zone is designed for shifting in irregular shape of mix-zone, which is more suitable to be deployed in this environment.

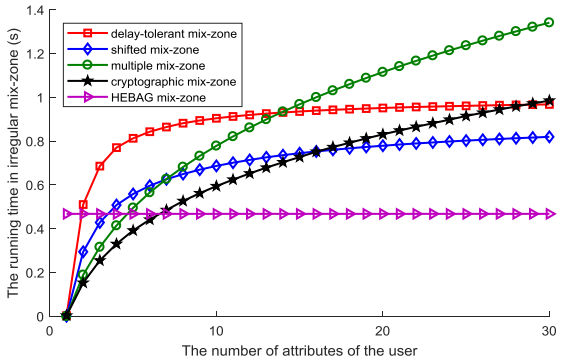


FIGURE 11. The running time in irregular mix-zone changed with the number of attributes (the number of users is 10).

From figure 12, we can see that the running time changed with the number of users is much more different to the running time changed with the number of attributes. Firstly, the running time of HEBAG mix-zone does not remain unchanged, but ascending with the increasing of user number.

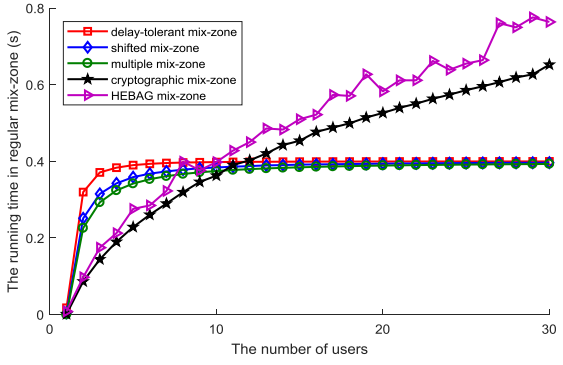


FIGURE 12. The running time in regular mix-zone changed with the number of users (the number of attributes is 15).

This is because of that this algorithm has to execute security multi-party computation to deal with attribute generalization, the increasing number of users will increase the number of individuals that calculates the similar attributes, so the running time becomes increasingly. Secondly, the cryptographic mix-zone also affected by the execution of encrypting attributes information and the running time also ascending with the increasing number of users. At last, the other algorithms do not sharply changed with the increasing number of users, as these algorithms were mainly designed to cope with the increasing number of attributes but not the user, so the running time does not be affected by the increasing number of users.

Similar to the running time in regular mix-zone, the running time in irregular mix-zone changed with the number of users is also affected seriously by the increasing number of users. From figure 13, the running time of HEBAG mix-zone still higher than others when the number of users increased, and the running time of cryptographic mix-zone also affected by the execution of encrypting attributes information. But the affection of irregular mix-zone for multiple mix-zones is descending, as this algorithm does not have to deploy more zones in irregular mix-zone than deployed in regular mix-zone, and the running time for disposing attribute generalization is descending. However, the tendency of descending is also restricted by the procedure of disposing attribute generalization.

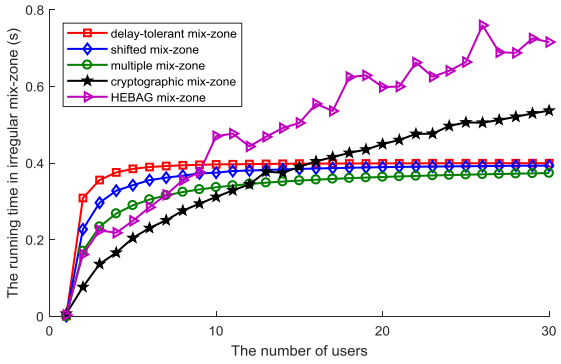


FIGURE 13. The running time in irregular mix-zone changed with the number of users (the number of attributes is 15).

Figure 14 shows the success ratio of all algorithms. In this figure, the success ratios of all algorithms are descending with the increasing number of users in regular mix-zone, as more users mean a higher anonymous value and the mix-zone has to involve more users to satisfy k-anonymity, so the algorithm will be failed when does not find enough users in current mix-zone. Among these algorithms, the affection for HEBAG mix-zone is the lowest, as this algorithm just needs to find enough number of users to execute the security multi-party computation to deal with attribute generalization, it does not need to select users with the restriction of similar attributes, so the success ratio is higher than other algorithms. The success ratio of multiple mix-zones is just a bit lower than

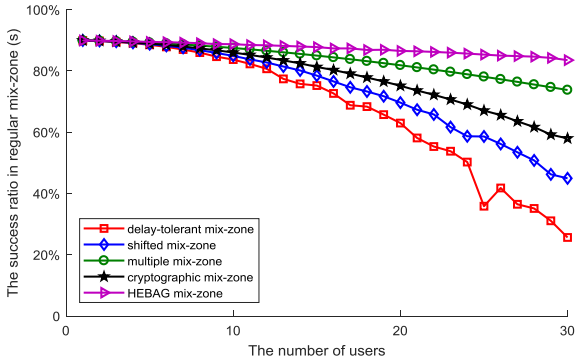


FIGURE 14. The success ratio in regular mix-zone changed with the number of users (the number of attributes is 15).

HEBAG mix-zone, as this algorithm can extend or repeatable deploy the zone to involve more users to ascend the success ratio. Similar with HEBAG mix-zone, cryptographic mix-zone also utilizes attribute encryption to protect the privacy, but this algorithm has to select users with similar attributes to achieve generalization and the procedure of user selection has affected the success ratio, as the user with similar attributes is difficult to be found and the insufficient of users with similar attributes will lead the failure of algorithm execution. The shifted mix-zone has improved the success ratio with region shifting, and its zone can shift to the region that involves more users with similar attributes, but this algorithm utilizes the delay tolerant to achieve query interval generalization and the delay tolerant has occupied more running times, which may lead some users cannot tolerate the delay time and reduce the success ratio. At last, the delay-tolerant mix-zone does not have the ability to shift to more users and also affected by the delay tolerant, so it success ratio is the lowest one.

The success ratio in irregular mix-zone changed with the number of users is similar to the success ratio in regular mix-zone, but the difference is that the success ratio is higher than regular mix-zone, so from figure 15 we can see that success ratios of all algorithms are ascending than these algorithms performed in regular mix-zone. The reason for the phenomenon can be ascribed as the irregular mix-zone involves more users than regular mix-zone, and increasing number of users has provided more users for disposing attribute generalization which then improves the success ratio.

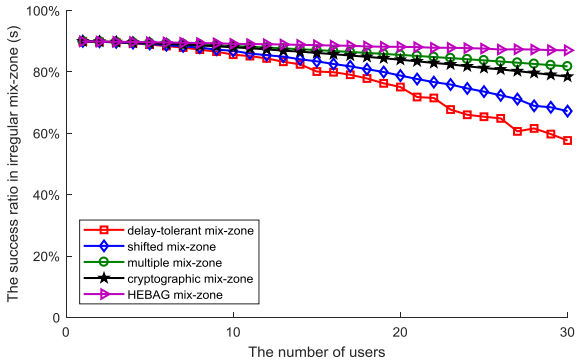


FIGURE 15. The success ratio in irregular mix-zone changed with the number of users (the number of attributes is 15).

Figure 16 shows the success ratio in regular mix-zone changed with the number of attributes. In this figure, the affection of success ratio of HEBAG mix-zone is minor, just descends when the number of attributes higher than 15. This is because of that the HEBAG mix-zone utilizes attribute generalization to protect the personal privacy, and the procedure of attribute generalization can deal with nearly all types of attributes in one round but not the attribute alone, so the success ratio does not be affected by the increasing number of attributes. For other algorithms, as these algorithms all designed to deal with attribute generalization by similar attributes selecting, and they also have to find users with similar attributes and need to deal with attributes by time delay and so on, so success ratios of these algorithms are lower than HEBAG mix-zone. In addition, as the disposed number of attributes is limited, the descending success ratios of multiple mix-zones, the delay-tolerant mix-zone and the shifted mix-zone are all present a tendency of smooth and steady. The reason for this phenomenon can be ascribed as these algorithms just dispose a specific number of attributes and then finish the execution, so the termination of execution can be seen as success executing these algorithms, but they failed to dispose other attributes and failed to protect the personal privacy.

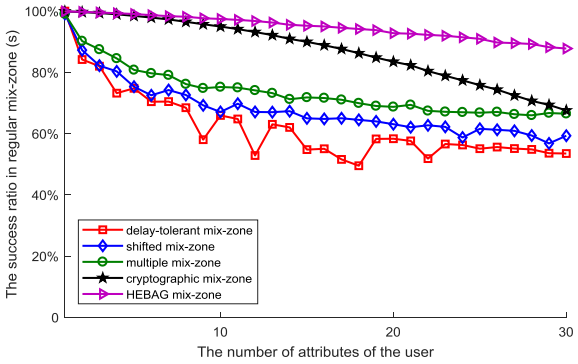


FIGURE 16. The success ratio in regular mix-zone changed with the number of attributes (the number of users is 10).

Figure 17 shows success ratios of different algorithms in irregular mix-zone changed with the number of attributes. Compared with the success ratio in regular mix-zone, we can see that the success ratio in irregular mix-zone is higher. The reason for this phenomenon can be ascribed as the irregular mix-zone involves more users than regular mix-zone, and the sufficient number of users improves the success ratio of finding users with similar attributes.

In conclusion, from above results as well as briefly explanation for each result, we can conclude that the HEBAG mix-zone is superior to other similar algorithms no matter in the performance in privacy protection or the performance in efficiency of algorithm execution, so it is better to be

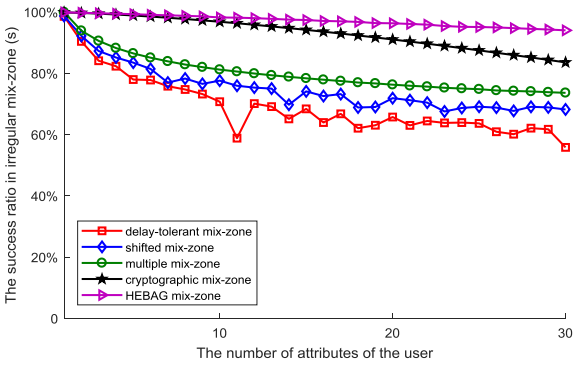


FIGURE 17. The success ratio in irregular mix-zone changed with the number of attributes (the number of users is 10).

deployed in real road network to provide service for personal privacy protection.

V. CONCLUSION

In the environment of road network, mix-zone is an efficient scheme used to resist attacks of tracking and attribute correlation. However, before deploying it into the practical environment, there still contains two problems unsolved. The first one is some attributes unnoticed will be utilized by the adversary to correlate the user. For another, some disguisers maybe pretended to be mix-zone users and be involved in the construction of mix-zone to grasp the private information of the user. In order to cope with above two problems, in this paper we propose a scheme called homomorphic encryption based attribute generalization mix-zone which based on the conception of attribute generalization as well as the homomorphic encryption. In this mix-zone, during the procedure of attribute generalization, all information transformed in the mix-zone is encrypted and no entities in the mix-zone can obtain any information about the user. Two phases of the attribute generalization are all secure, as the phase of bidding comparison to choose the agency all bidding values are encrypted, and the phase of calculation of similar attributes all information about attributes are encrypted. At last, in order to demonstrate the superiority of our proposed mix-zone, several analyses about the security and experiments are given, and then both of the effectiveness of privacy protection as well as the efficiency of algorithm execution are verified. Furthermore, the comparison results with other similar schemes further demonstrate the practicability of our proposed mix-zone. However, although our proposed mix-zone can solve the problem of attribute generalization as well as disguisers involved into the mix-zone, there still contains some other problems unsolved, such as the excess of calculation cost and the excess of disposing times, so the future work will focus on how to promote the efficiency of algorithm and so on.

REFERENCES

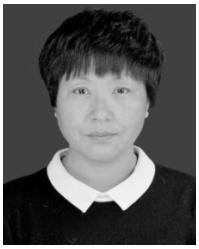
[1] X. Wang *et al.*, "A privacy-preserving message forwarding framework for opportunistic cloud of things," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 5281–5295, Dec. 2018.
[2] X. Wang *et al.*, "Privacy-preserving content dissemination for vehicular social networks: Challenges and solutions," *IEEE Commun. Surveys Tuts.*, to be published.
[3] M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking," in *Proc. 1st Int. Conf. Mobile Syst. Appl. Services*, May 2003, pp. 31–42.
[4] B. Gedik and L. Ling, "Location privacy in mobile systems: A personalized anonymization model," in *Proc. 25th IEEE Int. Conf. Distrib. Comput. Syst.*, Jun. 2005, pp. 620–629.
[5] B. Bamba, L. Liu, P. Pesti, and T. Wang, "Supporting anonymous location queries in mobile environments with privacygrid," in *Proc. 17th Int. Conf. World Wide Web*, Beijing, China, Apr. 2008, pp. 237–246.
[6] L. Fuyu, K. A. Hua, and C. Ying, "Query l-diversity in location-based services," in *Proc. 10th Int. Conf. Mobile Data Manage. Syst., Services Middleware*, May 2009, pp. 436–442.
[7] B. Niu, X. Y. Zhu, Q. H. Li, J. Chen, and H. Li, "A novel attack to spatial cloaking schemes in location-based services," *Future Gener. Comput. Syst.*, vol. 49, no. 2015, pp. 125–132, Aug. 2015.
[8] F. Fei, S. Li, H. Dai, C. Hu, W. Dou, and Q. Ni, "A K-anonymity based schema for location privacy preservation," *IEEE Trans. Sustain. Comput.*, to be published.
[9] Y. M. Sun, M. Chen, L. Hu, Y. F. Qian, and M. M. Hassan, "ASA: Against statistical attacks for privacy-aware users in location based service," *Future Gener. Comput. Syst.*, vol. 70, no. 2017, pp. 48–58, May 2017.
[10] 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.
[11] 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.
[12] Y. Wang, Y. Xia, J. Hou, S. M. Gao, X. Nie, and Q. Wang, "A fast privacy-preserving framework for continuous location-based queries in road networks," *J. Netw. Comput. Appl.*, vol. 53, pp. 57–73, Jul. 2015.
[13] K. Zeberga, R. Jin, H. J. Cho, and T. S. Chung, "A safe-region approach to a moving k-RNN queries in a directed road network," *J. Circuits Syst. Comput.*, vol. 26, no. 5, May 2017, Art. no. 1750071.
[14] M. Sepahkar and M. R. Khayyambashi, "A novel collaborative approach for location prediction in mobile networks," *Wireless Netw.*, vol. 24, no. 1, pp. 283–294, Jan. 2018.
[15] 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, 2013.
[16] B. Palanisamy, S. Ravichandran, L. Liu, B. Han, K. Lee, and C. Pu,, "Road network mix-zones for anonymous location based services," in *Proc. IEEE 29th Int. Conf. Data Eng. (ICDE)*, Apr. 2013, pp. 1300–1303.
[17] B. Palanisamy and L. Liu, "Effective mix-zone anonymization techniques for mobile travelers," *Geoinformatica*, vol. 18, no. 1, pp. 135–164, 2014.
[18] J. W. Kang *et al.*, "Location privacy attacks and defenses in cloud-enabled Internet of vehicles," *IEEE Wireless Commun.*, vol. 23, no. 5, pp. 52–59, Oct. 2016.
[19] Q. A. Arain *et al.*, "Privacy preserving dynamic pseudonym-based multiple mix-zones authentication protocol over road networks," *Wireless Pers. Commun.*, vol. 95, no. 2, pp. 505–521, Jul. 2017.
[20] L. Zhang, "OTIBAAGKA: A new security tool for cryptographic mix-zone establishment in vehicular ad hoc networks," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 12, pp. 2998–3010, Dec. 2017.
[21] Z. Lei, M. Chunguang, Y. Songtao, and Z. Xiaodong, "Location privacy protection model and algorithm based on profiles generalization," *Syst. Eng. Electron.*, vol. 38, no. 12, pp. 2894–2900, 2016.
[22] Z. Lei, M. Chun-Guang, Y. Song-Tao, and L. Zeng-Peng, "CP-ABE based users collaborative privacy protection scheme for continuous query," *J. Commun.*, vol. 38, no. 09, pp. 76–85, Aug. 2017.
[23] L. Zhang, S. Yang, J. Li, and L. Yu, "A particle swarm optimization clustering-based attribute generalization privacy protection scheme," *J. Circuits, Syst. Comput.*, vol. 27, no. 11, pp. 1850179-1–1850179-21, Oct. 2018.
[24] T. Dargahi, M. Ambrosin, M. Conti, and N. Asokan, "ABAKA: A novel attribute-based k-anonymous collaborative solution for LBSs," *Comput. Commun.*, vol. 85, pp. 1–13, Jul. 2016.
[25] Z. Ning, F. Xia, N. Ullah, X. J. Kong, and X. P. Hu, "Vehicular social networks: Enabling smart mobility," *IEEE Commun. Mag.*, vol. 55, no. 5, pp. 16–55, May 2017.

[26] Z. Ning, X. Wang, and J. Huang, "Mobile edge computing-enabled 5G vehicular networks: Toward the integration of communication and computing," *IEEE Veh.Technol. Mag.*, vol. 14, no. 1, pp. 54–61, Mar. 2018.
[27] X. Wang *et al.*, "Optimizing content dissemination for real-time traffic management in large-scale Internet of vehicle systems," *IEEE Trans. Veh. Technol.*, vol. 68, np. 2, pp. 1093–1105, Feb. 2018.
[28] T. Quan-you, M. Chuan-gui, and G. Yan, "Comparing private numbers based on fully homomorphic encryption," *J. Inf. Eng. Univ.*, vol. 6, no. 13, pp. 654–663, 2012.
[29] X. F. Wang, Y. Mu, and R. M. Chen, "One-round privacy-preserving meeting location determination for smartphone applications," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 8, pp. 1723–1732, Aug. 2016.
[30] B. Palanisamy, L. Liu, K. Lee, S. Meng, Y. Tang, and Y. Zhou, "Anonymizing continuous queries with delay-tolerant mix-zones over road networks," *Distrib. Parallel Database*, vol. 32, no. 1, pp. 91–118, Mar. 2014.
[31] 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.

ZHANG LEI was born in 1982. He received the B.E. and M.E. degrees from the College of Information Science and Electronic Technology, Jiamusi University, Jiamusi, China, in 2005 and 2011, respectively, and the Ph.D. degree from the College of Computer Science and Technology, Harbin Engineering University, in 2018. He is currently an Associate Professor and the Director of the Computer Software Teaching and Research Office, College of Information Science and Electronic Technology, Jiamusi University. His research interests include security and privacy in vehicle networks, and mobile privacy protocol.

HE LILI was born in 1979. She received the B.E. degree from the College of Information Science and Electronic Technology, Jiamusi University, Jiamusi, China, in 2001, and the M.E. degree from the College of Software Engineering, Beijing University of Posts and Telecommunications, Beijing, China, in 2014. She is currently a Lecturer of computer software with Jiamusi University. Her research interest includes software engineering.

LIU DESHENG was born in 1979. He received the Ph.D. degree in mechanical design and theory from Northeast Forestry University, China. He is currently a Professor with Jiamusi University, China. His research interests include the Internet of Things, intelligent control, and machine learning.



LI JING was born in 1968. She is currently a Professor with the College of Information and Electronic Technology, Jiamusi University, Jiamusi, China. Her research interests include machine learning and data privacy.

JIANG QINGFENG was born in 1983. He received the M.S. and Ph.D. degrees in computer science and technology from Harbin Engineering University, in 2008 and 2017, respectively. He is currently a Lecturer with the Changshu Institute of Technology and a member of the China Computer Federation. His research interests include computer network and information security.

YUAN QI received the Ph.D. degree from Harbin Engineering University, China, in 2018. She is currently an Associate Professor with Qiqihar University, China. Her research interests include wireless sensor network security, game theory, and block chain technology.