PLOS ONE

RESEARCH ARTICLE

# A strategy to balance location privacy and positioning accuracy

**Li He*, Junqing Liu, Peiyao Du**

School of Computer Science and Technology, Chongqing University of Posts and Telecommunication, Chongqing, China

* heli@cqupt.edu.cn



OPEN ACCESS

**Citation:** He L, Liu J, Du P (2024) A strategy to balance location privacy and positioning accuracy. PLoS ONE 19(5): e0304446. https://doi.org/10.1371/journal.pone.0304446

**Editor:** Ayesha Maqbool, National University of Sciences and Technology NUST, PAKISTAN

**Received:** February 5, 2024

**Accepted:** May 13, 2024

**Published:** May 30, 2024



**Data Availability Statement:** All data set are available from the https://doi.org/10.6084/m9.figshare.25577268.v2 database.

**Funding:** This research was partly funded by the Chongqing Graduate Research Innovation Project (Grant No.CYS23432), the National Nature Science Foundation of China (Grant No.61602073). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

In privacy protection methods based on location services, constructing anonymous areas using location information shared by collaborative users is the main method. However, this collaborative process not only increases the risk of mobile users' location privacy being leaked, but also reduces positioning accuracy. In response to this problem, we propose a balancing strategy, which transforms the problem of protecting mobile users' location privacy and improving positioning accuracy into a balance issue between location privacy and positioning accuracy. The cooperation of mobile users with different collaborating users is then modeled as an objective optimization problem, and location privacy and positioning accuracy are evaluated separately to make different selection strategies. Finally, an optimization function is constructed to select the optimal selection strategies. Experimental results show that our proposed strategy can effectively achieve the balance between location privacy and positioning accuracy.

## Introduction

With the rapid development of the Internet of Things and 5G/6G networks, smart mobile devices have accelerated their upgrading [1, 2]. The popularity of smart mobile devices has also promoted the application of location-based services (LBS) [3]. LBS pertains to the utilization of geographic location coordinates and associated data derived from mobile devices, with the objective of furnishing users possessing such devices with information resources and foundational services pertinent to their specific locations. The scope of LBS applications encompasses cartographic utilities, points of interest retrieval, and navigation, among others. Early LBS systems were mainly used in military and civilian fields involving important national interests. At present, LBS has been widely used in situation awareness, traffic navigation, business services, leisure and entertainment, and other fields [4]. Users can obtain points of interest (POI) near the location, such as the nearest restaurants, supermarkets, hospitals, etc., by sending LBS service requests. It can be seen that location-based services have penetrated into all aspects of people's daily lives, and thoughtful and meticulous services make people increasingly inseparable from LBS.

However, in order to obtain necessary information services, users need to send their location information to the LBS server, and the server returns corresponding query results to the

user based on the uploaded location information. At the same time, because the LBS server has the characteristics of honest but curious [5], it can make inferences based on the private information data uploaded by users, which may cause security issues for users [6]. For example, a malicious attacker may use private data to analyze the locations where users often stay, and determine whether the locations where users often stay during a specific period of time are home addresses and work locations. If malware cannot be prevented from spreading on the LBS server, or the LBS server does not have intrusion detection methods to detect malicious traffic, then this information will be cracked by malicious attackers, and users will face serious security and property threats [7–9].

In order to protect user location privacy, users need to communicate and cooperate with surrounding cooperating users to send the location information of multiple users to the server. However, working with more users means a reduction in positioning accuracy, thus affecting the accuracy of user location-based services. Therefore, it is crucial to balance location privacy and positioning accuracy for mobile users.

**Example** Taking Fig 1 as an example, if the mobile user $u$ cooperates with $u_1$, $u_2$, $u_3$, $u_4$, and $u_5$ simultaneously, it can be determined that $u$ is within the area jointly covered by the $u_1$, $u_2$, $u_3$, $u_4$, and $u_5$ response areas. This area is referred to as the inferred area in this paper, which is used to measure the accuracy of the inferred positioning accuracy. The distance between the cooperative user and the mobile user indicates whether the actual positioning is accurate, that is, the average distance between cooperative users and the mobile user is used to measure the accuracy of actual positioning accuracy. The accuracy of user positioning is determined by both the inferred positioning accuracy and the actual positioning accuracy. The degree of location privacy protection for the mobile user is also the same, and cannot be simply measured by

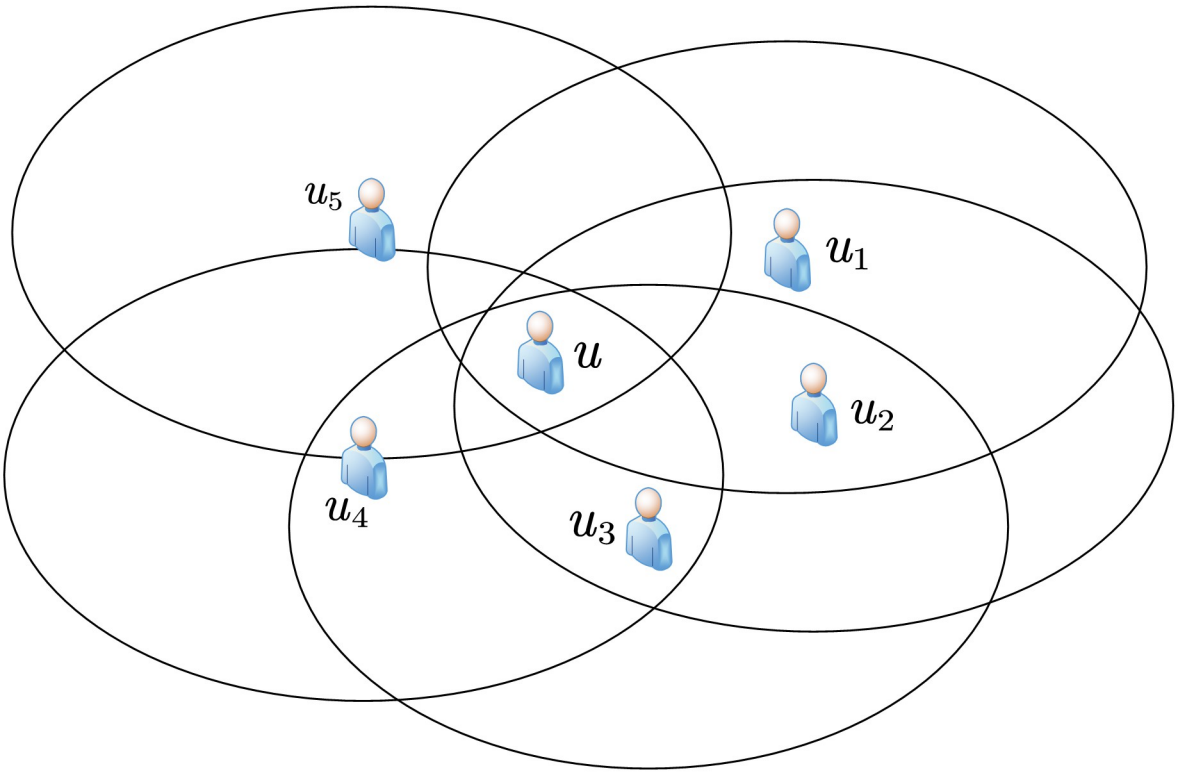


**Fig 1. Example of requesting user cooperation.**

https://doi.org/10.1371/journal.pone.0304446.g001

Yes or No. The convex hull area formed by $u_1$, $u_2$, $u_3$, $u_4$, and $u_5$ in Fig 1 is referred to as privacy area in this paper. The risk of the mobile user location information leakage depends on the size of the privacy area, because the smaller the privacy area, the closer the positions between cooperative users, thereby increasing the probability of mobile user location being discovered by attackers, and the higher the risk of location information leakage. In this paper, the key to protecting user privacy is to maximize the privacy area, as the larger the privacy area, the less likely the mobile user's location will be discovered by attackers. Therefore, cooperative users should stay as far away as possible from the mobile user. But this will reduce the actual positioning accuracy of user and improve their inferred positioning accuracy. In addition, the shorter the average distance between cooperative users and the mobile user, the higher the actual positioning accuracy of the mobile user. However, this increases the risk of user location privacy leakage. Reducing the actual positioning accuracy in turn leads to an increase in inferred positioning accuracy, which not only lowers the user's positioning accuracy but also increases the risk of location privacy leakage. Therefore, it is necessary to improve the actual positioning accuracy of users and reduce their inferred positioning accuracy while ensuring the privacy and security of user location. Therefore, we regard this issue as a balance among location privacy, actual positioning accuracy, and inferred positioning accuracy.

In response to the above problems, we propose a strategy to balance location privacy and positioning accuracy. The contributions of this paper are summarized as follows:

- We proposed a model to balance location privacy and positioning accuracy(BLPPA), and proved that it is NP-hard.
- We design a privacy evaluation model and a positioning accuracy evaluation model. It is proposed to use privacy areas to evaluate location privacy, while using actual distance and inferred areas to evaluate actual positioning accuracy and inferred positioning accuracy respectively. An optimization function is established to select the optimal decision to achieve a balance between location privacy and positioning accuracy.
- Based on the public real data set, experimental comparisons were conducted with existing solutions to verify the effectiveness of our designed model in balancing location privacy and positioning accuracy.

The remainder of this paper is organized as follows. Related work section introduces related work in the field of location privacy, and Problem statement section analyzes and formulates the BLPPA problem, proving the difficulty of the problem. Privacy and positioning accuracy assessment section evaluates location privacy and positioning accuracy respectively, and proposes an optimization function to find the optimal strategy. Experimental results section conducts experimental analysis. Conclusion section summarizes the full text.

## Related work

In order to solve the problem of location privacy leakage, researchers have proposed many algorithms, which are summarized as follows.

In the research on location privacy protection methods, anonymous methods are the more mature and most commonly used technology. Among the anonymous methods, *K*-anonymity [10] is one of the most classic algorithms.

*K*-anonymity is a data desensitization method. The core idea is to generalize quasi-identifiers so that any piece of data cannot be distinguished from at least $K-1$ other pieces of data. Gruteser et al. [11] first introduced the *K*-anonymity concept of relational databases into the field of LBS privacy protection and proposed location *K*-anonymity. It means that the position

in the LBS query corresponds to an area containing at least $K$ different users, and the attacker cannot distinguish the real query user from these $K$ users. A larger $K$ value indicates a higher degree of privacy protection, but the query accuracy will be lower, resulting in lower quality of service (QoS).

In recent years, researchers have been continuously proposing improvements and optimizations to $K$-anonymity techniques. Xing et al. [12] proposed a location privacy protection method based on double $K$-anonymity, which introduces a cloud server as a trusted third-party to isolate the direct communication between the user and the service provider, and at the same time reduces the relevance of the identity to the request through the method of substitution and combination, in order to hide the user's location and request information. Peng et al. [13] proposed a multidimensional privacy protection scheme that provides comprehensive protection for user privacy without the need for a trusted third party. The scheme employs a semi-trusted intermediate entity to perform user anonymization and blind filtering of results, utilizes Hilbert curves to transform the user's location, and uses encryption to preserve the user's query.

However, the above methods all rely on a third-party central server, which may store the user's real location information, and all queries submitted must go through it, which itself may become a bottleneck for system service performance and failure. In order to solve the shortcomings of centralized architecture, researchers have proposed distributed architecture [14–16]. Cui et al. [14] proposed a novel architecture that builds a non-localized LBS based on a distributed architecture, allowing mobile users to access the LBS without revealing their location. Furthermore, a technique to evaluate mobile user privacy and utility is proposed to achieve a balance between them. Shi et al. [15] considered that cooperative nodes would incur privacy costs when reporting their location information, and proposed a feasible incentive mechanism based on contract theory to reward cooperative nodes and ensure the expected positioning accuracy of the target node. Zhang et al. [16] proposed a cache-based double $K$-anonymity location privacy protection scheme, which reduces the load on user devices by applying multi-level caching and protects location privacy through double anonymity.

In order to address the issue of untrustworthy collaborators in distributed architectures, researchers have proposed conducting trust assessments on cooperative users to identify malicious users who disguise themselves as normal users and may engage in malicious operations [5, 17–20]. Luo et al. [17] and Li et al. [5, 18] proposed a blockchain based trust location privacy protection scheme in VANET. This scheme designs a trust management method based on Dirichlet distribution by analyzing the different requirements of requesting and cooperative vehicles in the process of constructing anonymous camouflage areas, so that requesting and cooperative vehicles only cooperate with the vehicles they trust. Liu et al. [19] proposed a blockchain-based TM scheme together with a conditional privacy-preserving announcement protocol (named as BTCPS). By the use of group signatures in anonymous aggregate vehicular announcement protocol, the reliability of announcements can be maintained without revealing users' privacy in the non-fully-trusted environment. Feng et al. [20] proposed a trusted CAC scheme called TCAC to protect the location privacy of vehicles. With the trust mechanism, multiple anonymizers in adjacent vehicular regions can be selected to construct the cloaking area in a cross-region manner. Min et al. [21] proposed a location privacy protection method in 3D space based on geo-indistinguishability, which develop a mechanism of three-variates Laplacian to generate perturbed locations considering the locations' X, Y, and Z-coordinates simultaneously, guaranteeing geo-indistinguishability. Furthermore, the truncation of the Laplace mechanism was further studied to limit the generated perturbation locations to specific regions. Kim et al. [22] used the perturbation mechanism of Geo-I to obfuscate user location information, and then proposed an expectation-maximization (EM) algorithm and

the deep learning based approaches to accurately calculate the density distribution of LBS users while preserving the privacy of location datasets.

In the field of location privacy protection, differential privacy [23] is also a popular research direction, which is characterized by being unaffected by attackers with background knowledge. Andrés et al. [24] proposed Geo-indistinguishability, where noise is added to the user's true location so that the final published user's location is within a circular range of the true location, making it impossible for the service provider to get the user's true geographic location from the collected location information. Li et al. [25] proposed an enhanced privacy definition beyond Geo-indistinguishability, combining with differential private indexing mechanism to design a new mechanism to realize this definition, by guaranteeing that the user's pseudo-location is reasonable to prevent the user's location perturbation behavior from being recognized. Wang et al. [26] proposed a location privacy preserving algorithm with location clustering and differential privacy, which firstly divides the continuous locations into different clusters, and then adds Laplacian noise to the stationary points and centers of mass within the clusters to protect the user's location privacy.

In the field of industrial Internet of Things, some research on privacy protection can also bring some inspiration. Wu et al. [27] proposed a privacy-preserving offloading scheme based on stochastic game theory considering multiple access points. In terms of privacy, the privacy risks caused by the offloading preferences of different edge nodes are studied, and the privacy entropy is used to evaluate the privacy protection level. Shen et al. [28] proposed a privacy protection model based on signaling game. This paper first derives the optimal privacy protection strategy of the model from a theoretical perspective. And a signaling Q-learning algorithm is designed to formulate the optimal privacy protection strategy by combining the Bayesian rule and the Q-learning approach. Wu et al. [29] first define the cumulative privacy amount for each IIoT user and trigger the privacy protection mechanism when the cumulative privacy amount exceeds the set privacy threshold. The offloading data generated by the IIoT user is then transferred to local processing, and finally, the cumulative privacy amount of the IIoT user is reduced.

The comparison of location privacy protection algorithms is shown in Table 1. Through the analysis of current research results, it can be found that the sharing of location information increases the risk of location privacy leakage for mobile users. At the same time, it also reduces the positioning accuracy. Therefore, we conduct research on this issue and propose solutions.

**Table 1. Comparison of location privacy protection algorithms.**

| References | Methods | Degree of location privacy protection | Accurate level of positioning accuracy |
|---|---|---|---|
| [12] | double k-anonymity | medium; no guarantee that cloud servers are completely trustworthy | not considered |
| [13] | encryption | medium; semi-trusted middle entity risk privacy leakage | not considered |
| [14] | a novel architecture | high; depend on the number of servers accessed simultaneously | medium; depending on privacy zone size |
| [15] | differential privacy | high; depend on how much noise is added | medium; depending on the degree of incentive of the incentive mechanism |
| [16] | double k-anonymity | high; depend on cache size | not considered |
| [5, 17–20] | blockchain | high; depend on the trust value threshold | not considered |
| [21–26] | differential privacy | high; depend on cache size | not considered |
| ours | k-anonymity | high; depend on the size of the privacy area, it can balance positioning accuracy while also protecting location privacy to a high degree. | high; using actual positioning accuracy and inferred positioning accuracy to comprehensively evaluate the positioning accuracy of mobile users |

https://doi.org/10.1371/journal.pone.0304446.t001

## Problem statement

### Preliminary

**Convex hull [30].** The minimum convex polygon encompassing all the points in a given point set Q is referred to as the convex hull. Since the convex hull problem investigates how to construct the smallest convex polygon that can enclose the given point set, we employ convex hull computation to determine the privacy area for cooperative users.

**3-SAT problem [31].** The 3-SAT problem is a Boolean satisfiability problem. Given a Boolean expression, it is necessary to find a variable assignment that makes the expression result true. If there exist assignments of variables as true or false that result in the Boolean expression evaluating to true, then the expression is satisfiable. If no such assignments exist, and for all possible variable assignments, the expression always evaluates to false, then the expression is unsatisfiable.

**Haversine formula [32].** The Haversine formula is a method for calculating the distance between two points on a great circle on the Earth's surface based on their longitudes and latitudes. It approximates the Earth as a sphere with a radius of R. This formula allows for the calculation of the great circle distance between any two points, A and B, on the Earth.

**Heron's formula [33].** The Heron's formula is a mathematical formula used to directly calculate the area of a triangle based on the lengths of its three sides.

### Problem analysis

When a mobile user has multiple cooperating users to cooperate with, a selection strategy needs to be made to determine which cooperating users the mobile user cooperates with to build an anonymous area. Taking the mobile user $u$ in the user layer in Fig 2 as an example,

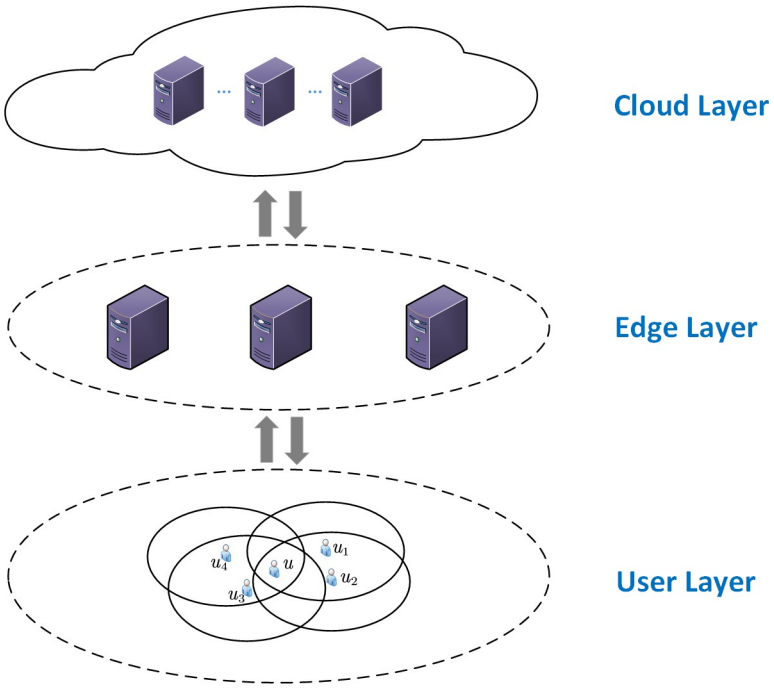


**Fig 2. System architecture.**

https://doi.org/10.1371/journal.pone.0304446.g002

**Table 2. Key notations.**

| Notation | Description |
|---|---|
| $u$ | the mobile user $u$ |
| $m$ | number of cooperative users |
| $N(u) = \{u_1, \ldots, u_i, \ldots, u_m\}$ | set of $u'$ cooperative users |
| $u_i \in N(u)$ | cooperative user $u_i$ |
| $a = (u_1, \ldots, u_i, \ldots, u_k)$ | selection strategy $a$ |
| $\boldsymbol{a} = \{a_1, \ldots, a_i, \ldots, a_{2^m}\}$ | BLPPA strategy $\boldsymbol{a}$ |
| $ad(a)$ | actual distance produced by $a$ |
| $ia(a)$ | inferred area produced by $a$ |
| $pa(a)$ | privacy area produced by $a$ |
| $\lambda_{ad}$ | weight of actual distance |
| $\lambda_{sa}$ | weight of inferred area |
| $\lambda_{pa}$ | weight of privacy area |
| $p_i$ | point $p_i$ |
| $circles(u_j)$ | coverage of $u_j$ |
| $d(u_i, u_j)$ or $d(p_i, p_j)$ | distance between $u_i$ or $p_i$ and $u_j$ or $p_j$ |
| $r$ | coverage radius of $u$ |
| $R$ | earth radius |
| $S(\Delta p_i p_j p_k)$ | triangle area formed by points $p_i$,$p_j$ and $p_k$ |
| $\widehat{S}(p_i O_{p_i p_j} p_j)$ | sector area formed by points $p_i$,$p_j$ and $O_{p_i p_j}$ |
| $\widehat{S}(p_i p_j)$ | arched area formed by points $p_i$,$p_j$ |

https://doi.org/10.1371/journal.pone.0304446.t002

there maybe be sixteen different selection strategies, namely $\{\{\oslash\}, \{u_1\}, \ldots, \{u_3\}, \{u_1, u_2\}, \ldots, \{u_2, u_3\}, \{u_1, u_2, u_3\}, \ldots, \{u_1, u_2, u_3, u_4\}\}$. These selection strategies result in different positioning accuracy and privacy area. For example, the selection strategy $\{u_1, u_2, u_3, u_4\}$ indicates that $u$ cooperates with all surrounding cooperative users at the same time, thus producing the largest privacy area. However, this selection strategy produces the highest inferred positioning accuracy. Too high inferred positioning accuracy will lead to too small inferred area and attacks. Combined with surrounding buildings and other information, the real location of the mobile user can be easily inferred, and this selection strategy produces the lowest actual positioning accuracy. In order to reduce the risk of the location being inferred and to improve the actual localization accuracy, $u$ will work with as few and as close collaborating users as possible to increase the inferred area size and shorten the average distance, but this leads to a reduction of the privacy area, making the anonymous area construction less effective and not conducive to privacy protection. Therefore, the solution to the BLPPA problem must achieve a balance among privacy area, actual positioning accuracy, inferred positioning accuracy generated for $u$. The main symbols and their interpretations in this paper are shown in Table 2.

## Problem formlation

**Definition 1 (Selection Strategy).** Given the mobile user $u$ and the set of cooperative users $u_i \in N(u)$, a selection strategy represents the collaboration of the mobile user $u$ with any cooperative user $u_i$. If $u$ collaborates with $k$ cooperative users, the selection strategy can be denoted as $a = (u_1, \ldots, u_i, \ldots, u_k)$, $i = 1, 2, \ldots, k; k \leq m$.

**Definition 2 (BLPPA Strategy).** The $u$'s BLPPA strategy is composed of a set of selection strategies represented by $\mathbf{a} = \{a_1, \ldots, a_i, \ldots, a_{2^m}\}$, $i = 1, 2, \ldots, 2^m$.

**Definition 3 (Average Distance).** Given the mobile user $u$ and the selection strategy $a = (u_1, \ldots, u_i, \ldots, u_k)$, $i = 1, 2, \ldots, k; k \leq m$, the average distance between $u$ and the cooperative users in $a$ is defined as $\overline{dis}(u, a)$.

**Definition 4 (Coverage).** Given the mobile user $u$, the request-response area of $u$ is defined as $cover(u)$.

**Definition 5 (Convex Hull Area).** Given the set of users $users = \{u_1, \ldots, u_i, \ldots, u_k\}$, $i = 1, 2, \ldots, k; k \leq m$, the convex hull area of the user set is defined as $cha(u_1, \ldots, u_i, \ldots, u_k)$.

The actual positioning accuracy generated by a selection strategy $a$ is measured by the length of the corresponding actual distance, defined as follows:

**Definition 6 (Actual Distance).** Given the selection strategy $a = (u_1, \ldots, u_i, \ldots, u_k)$, $i = 1, 2, \ldots, k; k \leq m$, the actual distance is the average distance between the cooperative users and the mobile user $u$, denoted as $ad(a)$, defined as follows:

$$ad(a) = \overline{dis}(u, a) \tag{1}$$

The inferred positioning accuracy generated by a selection strategy $a$ is measured by the size of the corresponding inferred area, defined as follows:

**Definition 7 (Inferred Area).** Given the selection strategy $a = (u_1, \ldots, u_i, \ldots, u_k)$, $i = 1, 2, \ldots, k; k \leq m$, the inferred area is the intersection of the coverage of the cooperative users, denoted as $ia(a)$, defined as follows:

$$ia(a) = \bigcap_{i=1}^{k} cover(u_i) \tag{2}$$

The privacy generated by a selection strategy $a$ is measured by the size of the corresponding privacy area, defined as follows:

**Definition 8 (Privacy Area).** Given the selection strategy $a = (u_1, \ldots, u_i, \ldots, u_k)$, $i = 1, 2, \ldots, k; k \leq m$, the privacy area is the convex hull area of the cooperative users, denoted as $pa(a)$, defined as follows:

$$pa(a) = cha(u_1, \ldots, u_i, \ldots, u_k) \tag{3}$$

## Problem hardness

In order to establish the NP-hardness of the BLPPA problem, it is necessary to reduce this problem to an NPC problem. Given that the 3-SAT problem is a known NPC problem, demonstrating that the BLPPA problem can be reduced to the 3-SAT problem will suffice to prove that the BLPPA problem is NP-hard.

**Theorem The BLPPA problem is NP-hard**

**Proof** First, it is necessary to formalize the problem in this paper, with three parameters: *ad*, *pa*, and *ia*, corresponding to actual distance, privacy area, and inferred area in this paper. If ↑ means increase, ↓ means decrease, the simplified constraint relationships among them are as follows:

$ad \uparrow \Rightarrow pa \uparrow, ia \downarrow$, meaning an increase in *ad* will lead to an increase in *pa* and a decrease in *ia*.

$pa \uparrow \Rightarrow ad \uparrow, ia \downarrow$, meaning an increase in *pa* will lead to an increase in *ad* and a decrease in *ia*.

$ia \uparrow \Rightarrow pa \downarrow, ad \downarrow$, meaning an increase in *ia* will lead to a decrease in *pa* and an increase in *ad*.

The ultimate goal is to find values for *ad*, *pa*, and *ia* that satisfy these constraint conditions, with the aim of making *ad* as small as possible to improve actual positioning accuracy, maximizing *ia* to reduce inferred positioning accuracy, and maximizing *pa* to enhance privacy.

For this purpose, it is proposed to create a 3-SAT problem that is equivalent to the problem in this paper. Given that an increase is true and a decrease is false, for each variable *ad*, *pa*, *ia*, introduce corresponding Boolean variables *ad_tf*, *pa_tf*, *ia_tf*. The constructed Boolean expression reflecting the constraint conditions is as follows:

$$(ad\_tf \vee pa\_tf \vee \overline{ia\_tf}) \wedge (ad\_tf \vee pa\_tf \vee \overline{ia\_tf}) \wedge (\overline{ad\_tf} \vee \overline{pa\_tf} \vee ia\_tf)$$

The Boolean variables that satisfy the objectives are $ad\_tf = false$, $pa\_tf = true$, and $ia\_tf = true$. Substituting these Boolean variables into the aforementioned Boolean expression, the process is as follows:

$$(ad\_tf \vee pa\_tf \vee \overline{ia\_tf}) \wedge (ad\_tf \vee pa\_tf \vee \overline{ia\_tf}) \wedge (\overline{ad\_tf} \vee \overline{pa\_tf} \vee ia\_tf)$$

$$\Rightarrow (flase \vee true \vee true) \wedge (false \vee true \vee \overline{true}) \wedge (\overline{false} \vee \overline{true} \vee true)$$

$$\Rightarrow true \wedge true \wedge true$$

$$\Rightarrow true$$

The final result shows that the result of Boolean expression is true.

Therefore, it has been proved that the original problem can be reduced to a 3-SAT problem, and the original problem is NP-hard.

## Privacy and positioning accuracy assessment

This section evaluates privacy protection and positioning accuracy respectively.

### Privacy assessment model

Privacy can be measured by the size of the corresponding privacy area. We propose to use the convex hull area composed of cooperative users to quantify the privacy area. In Fig 3(a), mobile user $u$ has five cooperating users, namely $u_1$, $u_2$, $u_3$, $u_4$, and $u_5$, so a selection strategy of $u$ is represented as $a = (u_1, u_2, u_3, u_4, u_5)$. Observing Fig 3(a), the privacy area consists of the

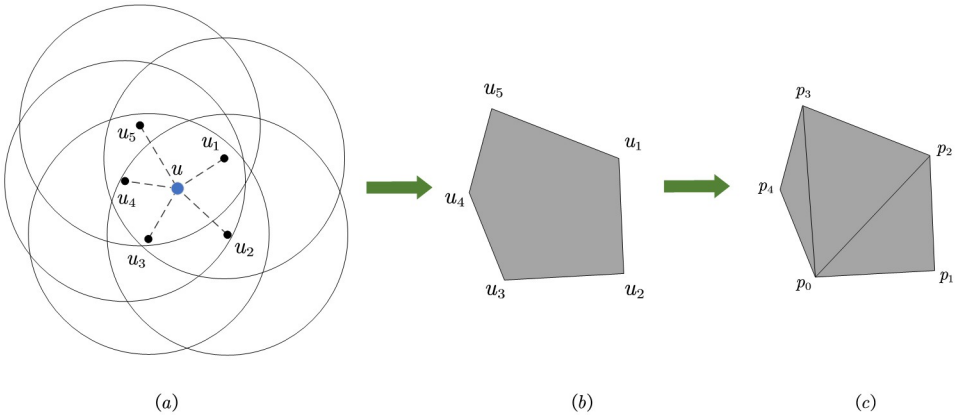


**Fig 3. Calculate the privacy area.**

https://doi.org/10.1371/journal.pone.0304446.g003

coordinate points of five cooperative users. The privacy area is extracted from Fig 3(a) to form the convex hull shown in Fig 3(b). Next, the area of this convex hull will be computed.

From Fig 3(b), it can be observed that the convex hull is an irregular polygon. By further subdividing the convex hull, it can be split into several triangles, as shown in Fig 3(c). The area of the convex hull can be calculated by summing the areas of the triangles that compose it. First, determine the set of points $P = \{p_2, p_1, p_0, p_4, p_3\}$ that make up the convex hull. The area of the convex hull formed by this point set is denoted as $cha(P)$. Select the lower-left point within the convex hull as the reference point, for example, use point $p_0$ in Fig 3(c) as the reference point. The remaining points are sorted based on the angles formed with the reference point and the positive x-axis. The sorted point coordinates might be represented as $\{p_1, p_2, p_3, p_4\}$. The sorted points and reference points form three triangles, and the sum of the areas of the three triangles is exactly equal to the convex hull area, i.e. $conv(P) = S(\Delta p_0p_1p_2) + S(\Delta p_0p_2p_3) + S(\Delta p_0p_3p_4)$. Below, we will take $S(\Delta p_0p_1p_2)$ as an example to discuss how to calculate the area of a triangle.

According to the Haversine Formula, the distance between $p_0$ and $p_1$ can be calculated based on their longitude and latitude $(lat_0, lng_0)$, $(lat_1, lng_1)$, and $(lat_2, lng_2)$, simply represented as $(x_0, y_0)$, $(x_1, y_1)$, and $(x_2, y_2)$:

$$d(p_0, p_1) = 2R \cdot \arcsin\left(\sqrt{\sin^2\frac{y_1 - y_0}{2} + \cos(y_0) \cdot \cos(y_1) \cdot \sin^2\frac{x_1 - x_0}{2}}\right) \tag{4}$$

where $R$ is the radius of the earth. In the same way, $d(p_0, p_2)$ and $d(p_1, p_2)$ can be calculated. According to Heron's Formula, $S(\Delta p_0p_1p_2)$ can be calculated as:

$$\begin{aligned} S(\Delta p_0p_1p_2) &= \sqrt{s(s - d(p_0, p_1))(s - d(p_0, p_2))(s - d(p_1, p_2))} \\ s &= \frac{d(p_0, p_1) + d(p_0, p_2) + d(p_1, p_2)}{2} \end{aligned} \tag{5}$$

Now, the formula for calculating the privacy area of mobile user $u$'s selection strategy $a$ is given:

$$pa(a) = S(\Delta p_0p_1p_2) + S(\Delta p_0p_2p_3) + S(\Delta p_0p_3p_4) \tag{6}$$

## Positioning accuracy assessment model

There are two types of positioning accuracy: actual positioning accuracy and inferred positioning accuracy. Firstly, the evaluation methods for actual positioning accuracy will be introduced, followed by the evaluation methods for inferred positioning accuracy.

**Actual positioning accuracy.** As introduced in Problem formlation section of this paper, the actual positioning accuracy is measured by the length of the corresponding actual distance, which is the geographic distance between the cooperative user $u_i$ and the mobile user $u$.

In Fig 4(a), mobile user $u$ has five cooperative users, namely $u_1$, $u_2$, $u_3$, $u_4$, and $u_5$. A selection strategy for $u$ is represented as $a = (u_1, u_2, u_3, u_4, u_5)$. The dashed line in Fig 4(a) indicates that the cooperative user has a cooperative relationship with the mobile user. Fig 4(b) is a simplified diagram in Fig 4(a), where all cooperative users and mobile users have quantified their cooperative relationships as actual distances, represented by solid lines. According to the Haversine Formula, the actual distance between mobile user $u$ and cooperative user $u_1$ can be calculated based on their latitude and longitude, which are simply denoted as $(x_u, y_u)$ and

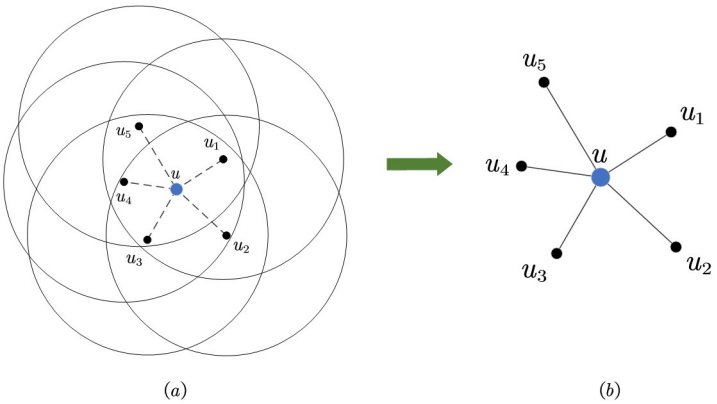


**Fig 4. Calculating the actual distance.**

https://doi.org/10.1371/journal.pone.0304446.g004

$(x_{u_1}, y_{u_1})$, and the calculation result is:

$$d(u, u_1) = 2R \cdot \arcsin\left(\sqrt{\sin^2 \frac{y_{u_1} - y_u}{2} + \cos(y_u) \cdot \cos\left(y_{u_1}\right) \cdot \sin^2 \frac{x_{u_1} - x_u}{2}}\right) \tag{7}$$

The actual distance calculation method between other cooperative users and the mobile user is similar. Finally, the formula for calculating the average actual distance of the mobile user $u$'s selection strategy $a$ is given:

$$ad(a) = \frac{\sum_{i=1}^{k} d(u, u_i)}{k} \tag{8}$$

**Inferred positioning accuracy.** The accuracy of inferred positioning is measured by the size of the corresponding inferred area, which is the intersection of the cooperative user coverage. In Fig 5(a), the mobile user $u$ has five cooperative users, namely $u_1$, $u_2$, $u_3$, $u_4$, and $u_5$. A same selection strategy for $u$ is represented as $a = (u_1, u_2, u_3, u_4, u_5)$. Extract the inferred area

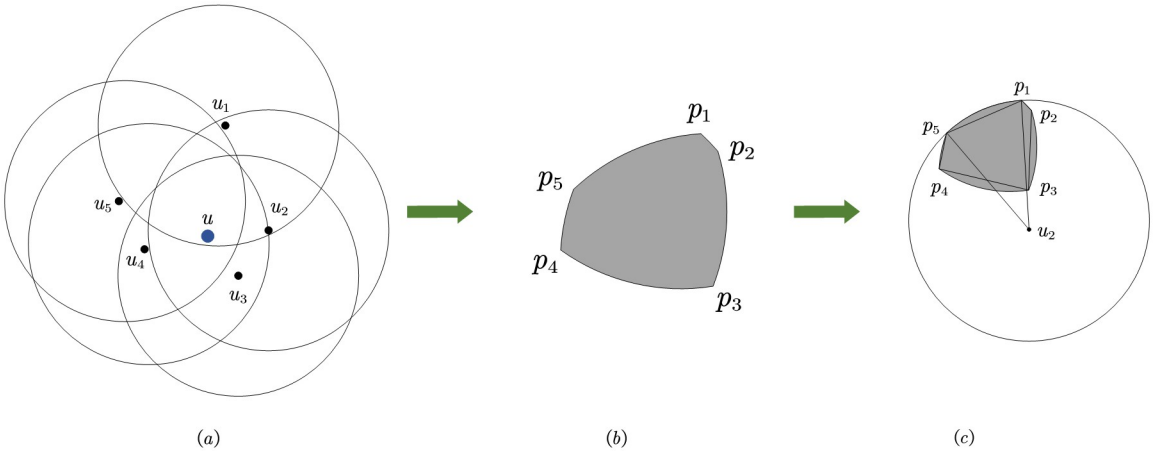


**Fig 5. Calculating the inferred area.**

https://doi.org/10.1371/journal.pone.0304446.g005

from Fig 5(a) to form the curved polygon shown in Fig 5(b). To calculate the area of the inferred area, the vertex of the curved polygon should be determined first. The method is as follows: calculate the intersection point of each cooperative user and other cooperative users in the requested area, determine whether each intersection point is an inner intersection point. If the distance between the intersection point and each cooperative user is not greater than the radius of the requested area, then the intersection point is considered an inner intersection point, which is a part of the curved polygon point set. Next, we will calculate the area of the curved polygon.

By observing Fig 5(b), it was found that the curved polygon can be divided into a convex hull and multiple arches, as shown in Fig 5(c). Therefore, the area of a curved polygon can be calculated by adding the areas of multiple arches of the convex hull at the center. The method for calculating the area of the convex hull has been detailed in Privacy Assessment Model section, so this section will discuss how to calculate the area of the arch.

Taking the calculation of the arch area $\widehat{S}(p_1p_5)$ as an example. Firstly, calculating the center angle formed by $p_1$, $p_5$, and $u_2$. If x represents longitude and y represents latitude, then the longitude and latitude of $p_1$, $p_5$, and $u_2$ are simply represented as $(x_1, y_1)$, $(x_5, y_5)$, and $(x_{u_2}, y_{u_2})$ respectively. The radius of the cooperative user coverage is $r$. Calculate the center angle $\theta(p_1u_2p_5)$ as:

$$\theta(p_1u_2p_5) = 2\arcsin\frac{d(p_1,p_5)}{2r} \tag{9}$$

Next, calculate the sector area $\widehat{S}(p_1u_2p_5)$ enclosed by $p_1$, $p_5$, and $u_2$:

$$\widehat{S}(p_1u_2p_5) = \pi r^2\frac{\theta(p_1u_2p_5)}{360} \tag{10}$$

The triangle area formed by $p_1$, $p_5$, and $u_2$ is $\widehat{S}(p_1u_2p_5)$, and the arch area $\widehat{S}(p_1p_5)$ can be calculated as:

$$\widehat{S}(p_1p_5) = \widehat{S}(p_1u_2p_5) - S(\Delta p_1u_2p_5) = r^2\cdot\arcsin\frac{d(p_1,p_5)}{2r} - S(\Delta p_1u_2p_5) \tag{11}$$

Finally, the privacy area calculation formula for $u$'s selection strategy $a$ is given:

$$ia(a) = pa(a) + \widehat{S}(p_1p_2) + \widehat{S}(p_2p_3) + \widehat{S}(p_3p_4) + \widehat{S}(p_4p_5) \tag{12}$$

## Algorithm analysis

It can be seen from Algorithm 1 that the time complexity of calculating the privacy area is $O(n)$. And $m$ cooperative users can generate $m - 2$ triangles. During the calculation process, the areas of $m - 2$ triangles need to be calculated one by one, and the time required is $O(n)$.

**Algorithm 1** Privacy Area Calculation

**Input:** The cooperative users $N(u) = \{u_1(x_1, y_1), u_2(x_2, y_2), \ldots u_n(x_n, y_n)\}$ for $u$, the selection strategy a, $R$

**Output:** Privacy area $pa(a)$

```
K ← 0, list_arctan ← ∅, pa(a) ← 0;
for u_i ∈ {u_1, ..., u_n} do
  if y_i < y_k or (y_i = = y_k and x_i<x_k) then
    k = i;
  end if
end for
```

7: let $u_k$ be the reference point;
8: swap the positions of $u_1$ and $u_k$ in $N(u)$;
9: **for** $u_i \in \{u_2, u_3, \ldots, u_n\}$ **do**
10: calculate the tan value between $u_1$ and $u_i$: $\tan(u_1, u_i) = (y_i - y_1)/(x_i - x_1)$;
11: calculate the arctan value between $u_1$ and $u_i$: $\arctan(u_1, u_i)$;
12: $list\_arctan \leftarrow arctan(u_1, u_i)$;
13: sort the $list\_arctan$ by the value of arctan value from smallest to largest;
14: **end for**
15: **for** $p_i \in \{u_2, u_3, \ldots, u_n\}$ **do**
16: calculate the distances $d(u_1, u_i)$ between $u_1$ and $u_i$, $d(u_1, u_{i+1})$ between $u_1$ and $u_{i+1}$, and $d(u_i, u_{i+1})$ between $u_i$ and $u_{i+1}$ according to formula (8);
17: calculate $S(\triangle u_1 u_i u_{i+1})$ according to formula (6);
18: $pa \leftarrow pa + S(\triangle u_1 u_i u_{i+1})$;
19: **end for**
20: return $pa(a)$;

From Algorithm 2, it can be seen that the time complexity of calculating the actual distance is also $O(n)$. In the calculation process, the geographic distances of $m$ cooperative users and requesting users need to be calculated separately, and then the mean value is calculated, which requires a time overhead of $O(n)$.

**Algorithm 2** Actual Distance Calculation

**Input:** The cooperative users $N(u) = \{u_1(x_1, y_1), u_2(x_2, y_2), \ldots u_n(x_n, y_n)\}$ for mobile user $u(x, y)$, the selection strategy $a$, $R$
**Output:** Actual Distance $ad(a)$
1: $ad(a) \leftarrow 0$, $dis \leftarrow 0$;
2: **for** $u_i \in \{u_1, u_2, \ldots, u_n\}$ **do**
3: calculate the distances $d(u, u_i)$ between $u$ and $u_i$ according to formula (8);
4: $dis \leftarrow dis + d(u, u_i)$;
5: **end for**
6: calculate average distance $ad(a)$: $ad(a) = \overline{dis}(u, \{u_1, u_2, \ldots, u_k\}) = \frac{dis}{n}$;
7: return $ad(a)$;

It can be seen from Algorithm 3 that the time complexity of calculating the size of the inferred area is $O(kn^2)$. We assume that the coverage of all users is a circle. The analysis steps are as follows:

Step 1: The coverage of any two cooperative users will produce two intersection points. The time complexity required to calculate all intersection points of the coverage of $m$ cooperative users is $O(n^2)$.

Step 2: If the distance between the intersection point and all cooperative users is not greater than the radius, the intersection point is regarded as an internal intersection point, and the time complexity required to determine whether all intersection points are internal intersection points is $O(n^2)$.

Step 3: Calculate the convex hull area formed by all internal intersection points according to Algorithm 1, the required time complexity is $O(n)$.

Step 4: Two adjacent inner intersection points form an arch, and the time complexity required to calculate the areas of all arches is $O(n)$.

**Algorithm 3** Inferred Area Calculation

**Input:** The cooperative users $N(u) = \{p_0(x_0, y_0), p_1(x_1, y_1), \ldots p_n(x_n, y_n)\}$ for $u$, the coverage radius $r$ of the cooperative user, the selection strategy $a$, $R$
**Output:** Inferred Area $sa(a)$
1: determine all circles $\{circes(u), circes(u_1), \ldots, circes(u_n)\}$;

```
candidate ← ∅, result ← ∅, ⌢S(result) ← 0, sa(a) ← 0;
for circles(u_i) ∈ {circles(u_1), circles(u_2), ..., circles(u_n)} do
  for circles(u_j) ∈ {circles(u_{i+1}), ..., circles(u_n)} do
    calculate intersection points of circles(u_i) and circles(u_j):
    p_0,p_1;
    candidate ← {p_0, p_1};
  end for
end for
for p_i ∈ candidate do
  for circles(u_k) ∈ {circles(u_1), circles(u_2), ..., circles(u_n)} do
    calculate the distance between p and circles(u_k):d(p_i, circles
    (u_k));
    if d(p_i, circles(u_k)) ≤ r then
      result ← p_i;
    end if
  end for
end for
calculate the convex hull area area(result) of result by Algorithm
1;
for p_i ∈ result do
  calculate the central angle of p_i, p_{i+1}, O_{p_ip_{i+1}}:
  θ(p_iO_{p_ip_{i+1}}p_{i+1}) ← 2arcsin(d(p_i,p_{i+1})/2r);
  calculate the sector area of p_i, p_{i+1}, O_{p_ip_{i+1}}:
  Ŝ(p_iO_{p_ip_{i+1}}p_{i+1}) ← πr²(θ(p_iO_{p_ip_{i+1}}p_{i+1})/360);
  calculate S(△p_0p_ip_{i+1}) according to formula (6);
  calculate the arch area: Ŝ(p_ip_{i+1}) ← Ŝ(p_iO_{p_ip_{i+1}}p_{i+1}) − S(p_iO_{p_ip_{i+1}}p_{i+1});
  Ŝ(result) ← Ŝ(result) + Ŝ(p_ip_{i+1});
end for
sa(a) ← Ŝ(result) + area(result);
return sa(a);
```

## Optimization function

To achieve a better balance among privacy area, actual positioning accuracy, and inferred positioning accuracy, it is necessary to select collaborative users in the optimal selection strategy to construct an anonymous area. The formation of optimal selection strategy must satisfy the following conditions:

In order to maximize the actual positioning accuracy, it is necessary to minimize the average distance $ad(a)$ between each collaborative user and the requesting user. Given $k$ collaborative users for mobile user $u$, the optimization function can be described as:

$$\begin{aligned} ad(a) &= \min_{a \in \mathbf{a}} ad(a) \\ &= \min_{a \in \mathbf{a}} \frac{\sum_{i=1}^{k} d(u, u_i)}{k} \end{aligned} \tag{13}$$

where the BLPPA strategy, denoted as **a**, consists of several selection strategies $a$.

To maximize the reduction of inferred positioning accuracy, it is necessary to maximize the common coverage area $ia(a)$ among the collaborative users. Given $k$ collaborative users for

mobile user $u$, denoted as $\{u_1, u_2, \ldots, u_k\}$, the optimization function can be described as:

$$\begin{aligned} ia(a) &= \max_{a\in\mathbf{a}} ia(a) \\ &= \max_{a\in\mathbf{a}} \left( \sum_{i=1}^{k-2} S(\Delta u_0 u_i u_{i+1}) + \sum_{i}^{k-1} \widehat{S}(u_i u_{i+1}) \right) \end{aligned} \tag{14}$$

To maximize the privacy area, it is necessary to maximize the convex hull area $pa(a)$ formed by the collaborative users. Similarly, given $k$ collaborative users for mobile user $u$, the optimization function can be described as:

$$\begin{aligned} pa(a) &= \max_{a\in\mathbf{a}} pa(a) \\ &= \max_{a\in\mathbf{a}} \sum_{i=1}^{k-2} S(\Delta u_0 u_i u_{i+1}) \end{aligned} \tag{15}$$

If the actual distance is minimized, the inferred area is maximized, and the privacy area is maximized at the same time, the optimal strategy selection problem is transformed into a multi-objective optimization problem.

$$\min_{\mathbf{a}} \left\{ \lambda_{ad} \cdot \underset{a\in\mathbf{a}}{ad}(a) + \lambda_{ia} \cdot \underset{a\in\mathbf{a}}{ia}(a) + \lambda_{pa} \cdot \underset{a\in\mathbf{a}}{pa}(a) \right\} \tag{16}$$

To address the above multi-objective optimization problem, we transform the multi-objective optimization problem into a single objective optimization problem by assigning weights to each objective function according to the relationship between the objective functions. Specifically, the multiple objective functions are linearly combined into a Privacy Positioning Accuracy Weighted Average (PPAWA) optimization function, and the optimal selection strategy is sought by adjusting the weights of each objective. Combining Eqs (13)–(16), the privacy positioning accuracy weighted average optimization function can be described as:

$$ppawa(a) = 1 - \min_{\mathbf{a}} \left\{ \lambda_{ad} \cdot \underset{a\in\mathbf{a}}{ad}(a) + \lambda_{ia} \cdot \underset{a\in\mathbf{a}}{ia}(a) + \lambda_{pa} \cdot \underset{a\in\mathbf{a}}{pa}(a) \right\} \tag{17}$$

where $\lambda_{ad}$, $\lambda_{ia}$, and $\lambda_{pa}$ represent the weights assigned to the actual distance, inferred area, and privacy area, respectively, indicating their relative importance in the BLPPA problem.

This method of solving the optimal strategy for the BLPPA problem is hereinafter referred to as Balance Location Privacy and Positioning Accuracy-Actual distance and Inferred area and Privacy area(BLPPA-AIP).

## Experimental analysis

We conducted a large number of experiments to test the performance of BLPPA-AIP, and selected five methods for comparison, as shown in Table 3.

### Experimental setup and performance metrics

The performance of BLPPA-AIP is tested using the publicly available dataset from the Geolife project of Microsoft Research Asia. In order to simulate different scenarios, we set up three scenarios: sparse scenario, normal scenario, and dense scenario, and the number of users in different scenarios is set to 100, 200, and 300, respectively [34], with the number of cooperative users taking the value of 10, and the radius of the user's coverage is 250 meters. In each experiment, given the location of a mobile user, we use a different approach to formulate a selection

**Table 3. Methods and descriptions of comparisons.**

| Methods | Description |
|---|---|
| BLPPA-AIP | Our method attempts to find the optimal selection strategy |
| APA(Actual Positioning Accuracy) | Only considering improving actual positioning accuracy |
| IPA(Inferred Positioning Accuracy) | Only considering reducing the accuracy of inferred positioning |
| PA(Privacy Area) | Only consider maximizing privacy area |
| RANDOM | Randomly select $k$ cooperative users from $n$ to cooperate |
| GREEDY | Select all cooperative users to cooperate with |

https://doi.org/10.1371/journal.pone.0304446.t003

strategy for the mobile user. The weights of privacy, actual location accuracy, and inferred location accuracy are all set to 1/3, giving all three equal importance.

For effectiveness assessment, four metrics are used: PPAWA, actual positioning accuracy, inferred positioning accuracy, and privacy area. To assess the efficiency, experimental comparisons are made with other methods, which are analyzed in terms of computation time and delay distance, respectively.

## Experimental results

**Effectiveness.** Tables 4–6 compare the average of privacy positioning accuracy, actual positioning accuracy, inferred positioning accuracy, and privacy area of six methods (i.e., BLPPA-AIP, APA, IPA, PA, EANDOM, GREEDY). Overall, compared to the other five methods, BLPPA-AIP performs best: it has the highest average PPAWA, the second largest inferred area, the second shortest average distance, and the fourth largest privacy area. This indicates that the scheme achieves an appropriate balance between location privacy and positioning accuracy.

From the PPAWA column in Tables 4–6, it can be seen that BLPPA-AIP achieved the highest PPAWA values in all three scenarios. In Table 4, the average PPAWA values are 26.32%,

**Table 4. Dense scenarios.**

| Method | PPAWA | Actual Distance | Inferred Area | Privacy Area |
|---|---|---|---|---|
| BLPPA-AIP | **0.7138** | 81.8627 | 64457.1538 | 11418.17 |
| APA | 0.5259 | **74.5726** | 35434.4131 | 2369.17 |
| IPA | 0.544 | 121.0637 | **68397.4418** | 635.31 |
| PA | 0.4981 | 162.5725 | 2425.916 | **39633.1** |
| RANDOM | 0.5233 | 139.5602 | 2425.916 | **39633.1** |
| GREEDY | 0.524 | 153.4682 | 50463.9713 | 13804.49 |

https://doi.org/10.1371/journal.pone.0304446.t004

**Table 5. Ordinary scenarios.**

| Method | PPAWA | Actual Distance | Inferred Area | Privacy Area |
|---|---|---|---|---|
| BLPPA-AIP | **0.7542** | 118.519 | 47555.8488 | 30688.73 |
| APA | 0.5632 | **109.1796** | 33844.5029 | 10564.84 |
| IPA | 0.5697 | 176.1167 | **76717.4008** | 3791.02 |
| PA | 0.6023 | 155.5322 | 8047.4802 | **42536.65** |
| RANDOM | 0.615 | 147.5397 | 8047.4802 | **42536.65** |
| GREEDY | 0.5314 | 156.8365 | 5190.5412 | 35328.99 |

https://doi.org/10.1371/journal.pone.0304446.t005

**Table 6. Sparse scenarios.**

| Method | PPAWA | Actual Distance | Inferred Area | Privacy Area |
|---|---|---|---|---|
| BLPPA-AIP | **0.707** | 183.9408 | 31842.6171 | 19198.71 |
| APA | 0.5873 | **174.4127** | 21696.7215 | 9048.57 |
| IPA | 0.6395 | 203.0802 | **31981.3572** | 6456.25 |
| PA | 0.5966 | 226.5934 | 643.3908 | **108453.87** |
| RANDOM | 0.6107 | 214.7921 | 643.3908 | **108453.87** |
| GREEDY | 0.5894 | 224.0338 | 685.1998 | 105014.16 |

https://doi.org/10.1371/journal.pone.0304446.t006

23.79%, 30.22%, 26.69%, and 26.59% higher than APA, IPA, PA, RANDOM, and GREDDY, respectively. This indicates that BLPPA-AIP can best balance location privacy and positioning accuracy. Meanwhile, from the Actual Distance column in Tables 4–6, it can be seen that the average distance generated by BLPPA-AIP to measure actual positioning accuracy is the second shortest among all methods, only 9.78% higher than the APA. From the Inferred Area column, it can be seen that the inferred area generated by BLPPA-AIP for measuring the accuracy of inferred positioning is the second largest among all methods, only 5.76% lower than the IPA. From the Privacy Area column, it can be seen that the comparison of privacy area sizes shows that BLPPA-AIP is worse than most methods and 71.19% lower than the PA. This result is expected, as APA, IPA, and PA only consider actual positioning accuracy, reduced inferred positioning accuracy, and privacy areas, respectively. When it is necessary to balance actual positioning accuracy, reduce inferred positioning accuracy, and privacy areas, these three methods are not suitable.

Take the performance of PA in ordinary scenarios as an example. Although PA achieves the maximum privacy area (Table 5), this method achieves the smallest inferred area and the longest average distance. This indicates that maximizing the user's privacy area without considering actual and inferred positioning accuracy will put mobile users at risk of location privacy leakage and reduced positioning accuracy. Similarly, the balance between position privacy and positioning accuracy achieved by APA is also extremely uneven. As shown in Table 5, although APA achieved the shortest average distance of 102.39 m, its privacy area is the second smallest, with 10564.84 $m^2$. The reason is that the goal of APA is to maximize the actual positioning accuracy, and in this process, it is inevitable to overlook the consideration of inferred positioning accuracy and privacy areas. Similar to APA, although IPA achieved the maximum inferred area of 76717.40 $m^2$, it had the longest average distance and the smallest privacy area, with 176.1167 m and 3791.02 $m^2$, respectively. The reason is that the goal of IPA is to minimize the actual inferred positioning accuracy, without considering the actual positioning accuracy and privacy area. This indicates that although APA and IPA can ensure the positioning accuracy of mobile users, they can easily lead to the leakage of user location privacy.

From Tables 4–6, it can be seen that all six methods have achieved smaller average distances, smaller inference areas, and smaller privacy areas in dense scenarios. This indicates that in dense scenarios, protecting the location privacy of mobile users is relatively difficult. However, dense scenes have the most mobile users and environmental information, and mobile users can infer corresponding locations and points of interest based on this. Therefore, in such dense scenarios, BLPPA-AIP is the best method because it provides a very high level of positioning accuracy without sacrificing too much location privacy.

**Efficiency.** *1) Computing time*. Due to the NP-hardness of the BLPPA problem, given mobile users in a specific area, if there are many cooperative users to choose from in the area, the time complexity of finding the selection strategy is high. Assuming there are $m$ cooperative

**Table 7. Computing time.**

| Method | BLPPA-AIP | APA | IPA | PA | GREEDY | RANDOM |
|---|---|---|---|---|---|---|
| Time (ms) | 77.795 | 12.988 | 60.840 | 19.971 | 10.995 | 10.995 |

https://doi.org/10.1371/journal.pone.0304446.t007

users around $u$, then there are a total of $2^m$ possible selection strategies. To evaluate the efficiency of BLPPA-AIP, we will discuss the computational time required to find a solution to the BLPPA problem.

Table 7 shows the calculation time evaluation results. The main reason why BLPPA-AIP takes the most time is that it needs to calculate all possible $2^m$ selection strategies for $m$ cooperative users for a given mobile user $u$, and select the best selection strategy among them. And the result of calculation time 77.795 ms is acceptable.

APA, IPA and PA take more time than the RANDOM and GREEDY with the goal of minimizing the average distance, maximizing the inferred area and maximizing the privacy area, respectively. The reason is because APA, in pursuit of minimizing the distance, only needs to select a cooperative user closest to $u$. Similarly, to maximize the privacy area, PA can simply select all cooperative users. These two methods do not need to compute all possible selection strategies as BLPPA-AIP does. In contrast, IPA pursues maximizing the inferred area and needs to compute all possible selection strategies like BLPPA-AIP, but it takes less time than BLPPA-AIP because it considers fewer factors. The reason that the GREEDY takes less time is because it selects all cooperative users and performs the computation of only one selection strategy. The reason that the RANDOM takes less time is because the RANDOM takes less time because it randomly selects one selection strategy and also performs the computation of only one selection strategy.

Fig 6 illustrates the effect of different numbers of cooperative users on the computation time. It can be seen that the computation time increases exponentially as the value of $m$ increases, this is because the method BLPPA-AIP we proposed needs to compute all possible $2^m$ selection strategies for $m$ cooperative users for cooperation given a mobile user $u$ and select the best selection strategy among them. So when the value of $m$ increases, the number of selection strategies that may need to be computed grows exponentially, which greatly increases the amount of computation and also leads to an exponential increase in computation time. From Fig 6, it can be seen that when the number of required cooperative users is small (e.g., when $m$ is 3 to 7), the computation time is low, but fewer cooperative users produce weaker privacy protection. When the number of required cooperative users is large (e.g., when $m$ is greater than 10), the computation time is too high to be acceptable, even though the privacy protection is excellent. Therefore, the value of $m$ should be moderate, and all the experiments in this paper are done when $m$ is 10.

*2) Delay distance* Delay distance is the combined measure of delay and distance experienced by the user in the moving process, which is specifically described as the user's moving distance from issuing a cooperation request to balancing location privacy protection and positioning accuracy, which reflects the efficiency of the method used for balancing location privacy protection and positioning accuracy, and the shorter the delay distance is, it indicates that the method handles balancing location privacy protection and positioning accuracy more efficiently in the process of the user's moving.

We set three speed metrics: 5km/h, 30km/h and 60km/h, which are used to simulate the user's movement at slow, medium, and fast speeds, respectively [35]. Table 8 demonstrates the comparison results of the delay distance of each method at different speeds. The delay distances of BLPPA-AIP methods are higher than the comparison methods at different speeds,

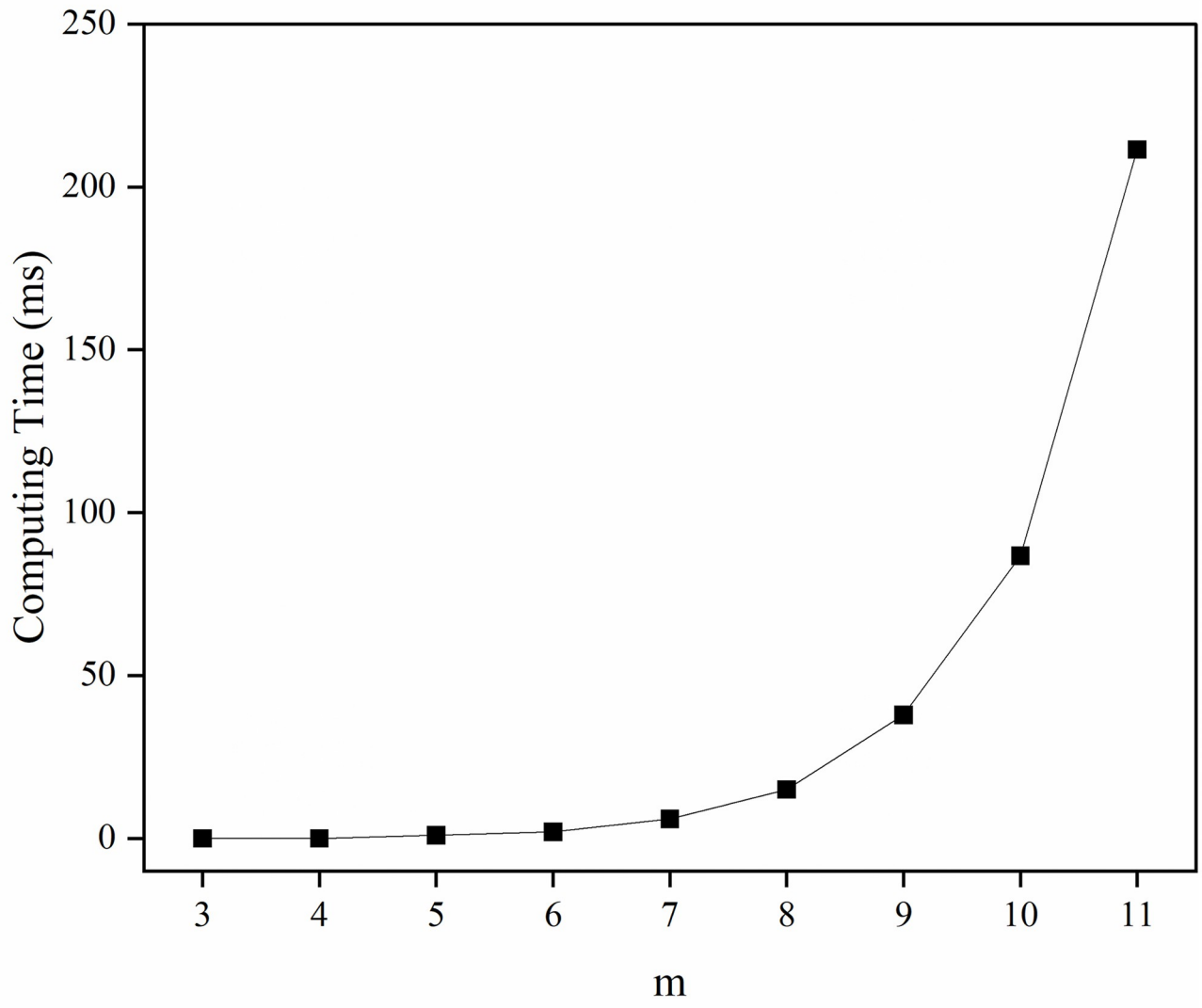


**Fig 6. Computing time for different *m*.**

https://doi.org/10.1371/journal.pone.0304446.g006

but the delay distances of BLPPA-AIP methods at different speeds are within the acceptable range.

When the user moves at a slow speed (e.g., walking), the delay distance of BLPPA-AIP is 0.15657 m, which is only 0.14 m lower than that of the GREEDY and RANDOM methods,

**Table 8. Comparison of delay distance for each method at different speeds (m).**

| Method | 5km/h | 30km/h | 60km/h |
|---|---|---|---|
| BLPPA-AIP | 0.15657 | 0.93942 | 1.87883 |
| APA | 0.01943 | 0.11656 | 0.23312 |
| IPA | 0.11363 | 0.68178 | 1.36355 |
| PA | 0.03324 | 0.19945 | 0.39890 |
| GREEDY | 0.01527 | 0.09164 | 0.18328 |
| RANDOM | 0.01527 | 0.09164 | 0.18328 |

https://doi.org/10.1371/journal.pone.0304446.t008

which have the shortest delay distances at the same speed, indicating that the delay distance of the BLPPA-AIP method does not affect the user's experience of the location-based service under the user's slow speed movement.

When the user moves at medium speed (e.g., riding), the delay distance of BLPPA-AIP is 0.93942 m, which is 0.85 m lower than that of GREEDY and RANDOM, which have the shortest delay distances at the same speed, indicating that under the user's medium-speed movement, the delay distance of the BLPPA-AIP method basically does not affect the user's experience of location-based services.

When the user is moving fast (e.g., driving), the delay distance of BLPPA-AIP is 1.87883 m, which is 1.69 m lower than that of the GREEDY and RANDOM methods that have the shortest delay distances at the same speed, indicating that under the user's fast movement, the delay distance of the BLPPA-AIP method basically does not affect the user's experience of the location-based service.

## Conclusion

In this paper, we propose a balanced location privacy and positioning accuracy strategy BLPPA-AIP. In terms of location privacy protection, BLPPA-AIP provides personalized location privacy protection for mobile users by constructing an anonymous area in the edge environment. In terms of improving location accuracy, BLPPA-AIP models the problem of selecting cooperative users as an objective optimization problem and constructs an optimization function to select cooperative users to ensure the highest location accuracy. In addition, we also propose methods to evaluate location privacy, actual positioning accuracy, and inferred positioning accuracy, so that BLPPA-AIP can achieve a balance between location privacy and positioning accuracy. Simulation results show that BLPPA-AIP not only achieves better location privacy, but also ensures high positioning accuracy, i.e., it strikes a proper balance between location privacy and positioning accuracy. In future work, how to further improve the efficiency as well as analyze the evaluation metrics affecting the choice of cooperative users at a finer granularity will be the next research focus.

## Author Contributions

**Conceptualization:** Junqing Liu.

**Data curation:** Junqing Liu, Peiyao Du.

**Investigation:** Peiyao Du.

**Methodology:** Junqing Liu.

**Supervision:** Li He.

**Validation:** Peiyao Du.

**Writing – original draft:** Junqing Liu.

**Writing – review & editing:** Li He, Junqing Liu.

## References

1. Shen Y., Shen S., Li Q., Zhou H., Wu Z., Qu Y. Evolutionary privacy-preserving learning strategies for edge-based iot data sharing schemes. Digital Communications and Networks, 9(4):906–919, 2023. https://doi.org/10.1016/j.dcan.2022.05.004
2. Wang C., Jiang C., Wang J., Shen S., Guo S., Zhang P. Blockchain-aided network resource orchestration in intelligent internet of things. IEEE Internet of Things Journal, 10(7):6151–6163, 2022. https://doi.org/10.1109/JIOT.2022.3222911

3. Ma C., Yan Z., Chen CW. SSPA-LBS: Scalable and social-friendly privacy-aware location-based services. IEEE Transactions on Multimedia, 21(8):2146–2156, 2019. https://doi.org/10.1109/TMM.2019.2892300
4. Jiang H., Zeng J., Han K, Liu Y. Research on location privacy protection methods for mobile users in 5g environment. Journal of Beijing Institute of Technology, 41(1):84–92, 2021.
5. Li B., Liang R., Zhou W., Yin H., Gao H., Cai K. LBS meets blockchain: An efficient method with security preserving trust in sagin. IEEE Internet of Things Journal, 9(8):5932–5942, 2022. https://doi.org/10.1109/JIOT.2021.3064357
6. Jiang H., Zhao P., Wang C. RobLoP: Towards robust privacy preserving against location dependent attacks in continuous lbs queries. IEEE/ACM Transactions on Networking, 26(2):1018–1032, 2018. https://doi.org/10.1109/TNET.2018.2812851
7. Yu S., Zhai R., Shen Y., Wu G., Zhang H., Yu S., et al. Deep q-network-based open-set intrusion detection solution for industrial internet of things. IEEE Internet of Things Journal, 2023. https://doi.org/10.1109/JIOT.2023.3333903
8. Shen S., Cai C., Li Z., Shen Y., Wu G., Yu S. Deep q-network-based heuristic intrusion detection against edge-based siot zero-day attacks. Applied Soft Computing, 150:111080, 2024. https://doi.org/10.1016/j.asoc.2023.111080
9. Shen S., Xie L., Zhang Y., Wu G., Zhang H., Yu S. Joint differential game and double deep q–networks for suppressing malware spread in industrial internet of things. IEEE Transactions on Information Forensics and Security, 2023. https://doi.org/10.1109/TIFS.2023.3307956
10. Sweeney L. k-anonymity: A model for protecting privacy. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 10(5):557–570, 2002. https://doi.org/10.1142/S0218488502001648
11. Gruteser M., Grunwald D. Anonymous usage of location-based services through spatial and temporal cloaking. In Proceedings of the 1st International Conference on Mobile Systems, Applications and Services, pages 31–42, 2003. https://doi.org/10.1145/1066116.1189037
12. Xing L., Jia X., Gao J., Wu H. A location privacy protection algorithm based on double k-anonymity in the social internet of vehicles. IEEE Communications Letters, 25(10):3199–3203, 2021. https://doi.org/10.1109/LCOMM.2021.3072671
13. Peng T., Liu Q., Wang G., Xiang Y., Chen SH. Multidimensional privacy preservation in location-based services. Future Generation Computer Systems, 93:312–326, 2019. https://doi.org/10.1016/j.future.2018.10.025
14. Cui G., He Q., Chen F., Jin H., Xiang Y., Yang Y. Location privacy protection via delocalization in 5g mobile edge computing environment. IEEE Transactions on Services Computing, 16(1):412–423, 2023. https://doi.org/10.1109/TSC.2021.3112659
15. Shi X., Yu D., Fu M., Zhang WA. Clap: A contract-based incentive mechanism for cooperative localization balancing localization accuracy and location privacy. IEEE Internet of Things Journal, 9(9):6678–6687, 2022. https://doi.org/10.1109/JIOT.2021.3110961
16. Zhang S., Hu B., Liang W., Li K., Gupta BB. A caching-based dual k-anonymous location privacy-preserving scheme for edge computing. IEEE Internet of Things Journal, 10(11):9768–9781, 2023. https://doi.org/10.1109/JIOT.2023.3235707
17. Luo B., Li X., Weng J., Guo J., Ma JF. Blockchain enabled trust-based location privacy protection scheme in vanet. IEEE Transactions on Vehicular Technology, 69(2):2034–2048, 2020. https://doi.org/10.1109/TVT.2019.2957744
18. Li B., Liang R., Zhu D., Chen W., Lin QY. Blockchain-based trust management model for location privacy preserving in vanet. IEEE Transactions on Intelligent Transportation Systems, 22(6):3765–3775, 2021. https://doi.org/10.1109/TITS.2020.3035869
19. Liu X., Huang H., Xiao F., Ma ZY. A blockchain-based trust management with conditional privacy-preserving announcement scheme for vanets. IEEE Internet of Things Journal, 7(5):4101–4112, 2019. https://doi.org/10.1109/JIOT.2019.2957421
20. Feng J., Wang Y., Wang J., Ren F. Blockchain-based data management and edge-assisted trusted cloaking area construction for location privacy protection in vehicular networks. IEEE Internet of Things Journal, 8(4):2087–2101, 2020. https://doi.org/10.1109/JIOT.2020.3038468
21. Min M., Xiao L., Ding J., Zhang H., Li S., Pan M., et al. 3d geo-indistinguishability for indoor location-based services. IEEE Transactions on Wireless Communications, 21(7):4682–4694, 2021. https://doi.org/10.1109/TWC.2021.3132464
22. Kim J., Lim B. Effective and privacy-preserving estimation of the density distribution of LBS users under geo-indistinguishability. Electronics, 12(4):917, 2023. https://doi.org/10.3390/electronics12040917
23. Dwork C. Differential privacy. In 33rd International Colloquium on Automata, Languages and Programming, volume 4052, pages 1–12, 2006. https://doi.org/10.1007/11787006_1

24. Andrés M., Bordenabe N., Chatzikokolakis K., Palamidessi C. Geo-indistinguishability: Differential privacy for location-based systems. In Proceedings of the 2013 ACM SIGSAC Conference on Computer and Communications Security, pages 901–914, 2013. https://doi.org/10.1145/2508859.2516735

25. Li X., Ren Y., Yang L., Zhang N., Luo B., Weng J., et al. Perturbation-hidden: Enhancement of vehicular privacy for location-based services in internet of vehicles. IEEE Transactions on Network Science and Engineering, 8(3):2073–2086, 2021. https://doi.org/10.1109/TNSE.2020.3011607

26. Wang B., Li H., Ren X., Guo Y. An efficient differential privacy-based method for location privacy protection in location-based services. Sensors, 23(11), 2023. https://doi.org/10.3390/s23115219 PMID: 37299946

27. Wu G., Chen X., Gao Z., Zhang H., Yu S., Shen SG. Privacy-preserving offloading scheme in multi-access mobile edge computing based on MADRL. Journal of Parallel and Distributed Computing, 183:104775, 2024. https://doi.org/10.1016/j.jpdc.2023.104775

28. Shen S., Wu X., Sun P., Zhou H., Wu Z., Yu S. Optimal privacy preservation strategies with signaling Q-learning for edge-computing-based IoT resource grant systems Expert Systems with Applications, 225:120192, 2023. https://doi.org/10.1016/j.eswa.2023.120192

29. Wu G., Chen X., Shen Y., Xu Z., Zhang H., Shen S., et al. Combining lyapunov optimization with actor-critic networks for privacy-aware IIoT computation offloading. IEEE Internet of Things Journal, 2024. https://doi.org/10.1109/JIOT.2024.3357110

30. Hazewinkel M. Encyclopaedia of mathematics. Kluwer Academic Publishers, 1990.

31. Tian NY. Study on several methods based on boolean satisfiability. Thesis, Jilin University, 2022.

32. Winarno E., Hadikurniawati W., Rosso RN. Location based service for presence system using haversine method. In 2017 International Conference on Innovative and Creative Information Technology (ICITech), pages 1–4, 2017. https://doi.org/10.1109/INNOCIT.2017.8319153

33. Weisstein, Eric W. Heron's formula. https://mathworld.wolfram.com/HeronsFormula.html, 2023.

34. Wu X. Research on privacy protection based on lbs service of mobile users. Thesis, Beijing Jiaotong University, 2022.

35. Yin ZZ. The behavior characteristics study of pedestrians, bicycles and vehicles and the construction of unified traffic model. Thesis, Fuzhou University, 2020.