EAI Endorsed Transactions
on e-Learning

Research Article EAI.EU

A K-Anonymous Location Privacy-Preserving Scheme for Mobile Terminals

Weiping Peng[1], Di Ma[1,*], Cheng Song[1], Daochen Cheng[1], Jiabao Liu[1,]

[1]School of Computer Science and Technology, Henan Polytechnic University, No. 2001, Century Road, Jiaozuo 454003, China

Abstract

Mobile terminals boost the prosperity of location-based service (LBS) which have already involved in every aspect of People's daily life and are increasingly used in various industries. Aimed at solving the security and efficiency problem in the existing location privacy protection schemes, a K-anonymity location privacy preservation scheme based on mobile terminal is proposed. Firstly, number of rational dummy locations is selected from the cloaking region, from which more favorable locations are further filtered according to location entropy to a better anonymity effect can be achieved. Secondly, the secure and efficient m-out-of-n oblivious transfer protocol is adopted, which not only avoids the dependency on the trusted anonymity center in existing schemes to improve the efficiency, but also meets the requirements for querying multiple interest points at one time. Security analyses demonstrate that the scheme satisfies such security properties as anonymity, non-forgeability and resistance to replay attack, and simulation results show that the scheme has higher execution efficiency and privacy level, while is low in communications costs.

Keywords: Location-based service, K-anonymity, Privacy protection, location terminals

Received on 24 November 2023, accepted on 6 December 2023, published on 11 December 2023

doi: 10.4108/eetel.4468

1. Introduction

Along with the development of computer technology, global positioning systems and wireless communication networks, location-based service (LBS) technology [1][2][3] has also increasingly popular. Location-based services, LBS, refers to a location service provider that obtains the location of a device through various types of positioning technology and provides the device with specific services requested by the device via the Internet. Its typical applications include vehicle navigation, online mailing, takeout and ticketing services, etc. It not only brings great convenience to users, but also causes some changes in people's daily behavior. However, as users become increasingly dependent on LBS, private location privacy may as well at a great risk of disclosure accordingly[4][5].

When applying for LBS, mobile user needs to submit its current location and query information to LBS server in real time, which may help to establish user's location trajectory based on the temporal-spatial relationship. Thereby, this user's location trajectory can be employed to infer user's whereabouts, family address, and work place, so as to further obtain such privacy information as religious belief, living habits, medical information[6][7], etc. If such privacy information is obtained by any attacker, then user's location-related privacy is at great risk of disclosure. Consequently, location privacy protection technology is one of the research focuses in current field of mobile network security.

Focusing on the problem of location privacy protection, this paper employs both K-anonymity technique and Oblivious Transfer (OT) protocol to propose a K-

* Corresponding author. Email: maddy@home.hpu.edu.cn

anonymity location privacy protection scheme based on mobile terminal (KBMT). In this scheme, user firstly sends to LBS server k number of ID (including that of real user) as registration request, and LBS server generates pseudonyms and public/private key pairs for k number of ID after receiving registration request. When registration is completed, mobile terminal will generate and select dummy locations to complete location service query on the basis of OT protocol and LBS server. This paper makes contributions as follows:

- We propose a K-anonymity location privacy protection scheme based on mobile terminal, which combines ID-based cryptosystem, K-anonymity technology and OT protocol to avoid privacy security's dependency on the trusted third party in the existing schemes, and to accomplish multiple services by one single request.
- We design a security enhancement algorithm for generating and selecting virtual locations to reduce the risk of privacy leakage of a user's real location by providing more confusing virtual locations.
- Our scheme verifies its effectiveness and security by establishing a simulation experiment environment.

The rest of this paper is organized as follows: In Sect. 2. Work related to the research presented in this paper is presented. Some necessary preliminaries are described in Sect. 3. The K-anonymity Privacy Protection Scheme based on Terminal is proposed in Sect. 4. The security analysis of the k-anonymous trajectory privacy protection scheme is given in Sect.5. Sect. 6 focuses on the simulation of the proposed scheme. And conclusions are drawn in Sect. 7.

2. Related Work

2.1. Location privacy

In recent years, some scholars home and abroad have done lots of research on privacy protection technology and achieved certain positive results[24][25]. According to the different architectures of location privacy protection systems, the results can be divided into two categories: location privacy protection technology based on the trusted anonymity center and location privacy protection technology based on mobile terminal.

Architecture based on the trusted anonymity center is also called trusted third-party (TTP) architecture, which is first proposed by Gedik and Liu[8] Location privacy protection scheme based on TTP architecture introduces the trusted anonymity center amid user and LBS server, and the center usually adopts some privacy protection techniques [9][10][11] to anonymize user's message of service request, as well as to complete the message transfer between user and LBS server, which in turn succeeds in protecting user's location privacy, and reducing the storage and computation costs of user's terminal. K-anonymity technology[12][13][14][15], as the most widely used location privacy protection technology, mainly forms a cloaking region including at least other number of different users and then this cloaking region replaces real user to send service request to LBS server, so as to reduce the precision of user's location, which will finally make the probability of attacker identifying the real user in cloaking area less than $1/K$. In order to protect the safety of both the user's location and the query location, Kuang et al. [16]carried out a bidirectional K- disturbance based on user's location and query location semantics, and the user matched the road section with the highest security according to the sensitivity preference and specified the K anonymity. K-anonymity can prevent the identity information leak, but fails to prevent the attribute information leak. To solve this problem, Tu et al. [17]proposed a privacy preserving scheme to prevent semantic and re-identification attacks by employing three data masking methods: k-anonymity, l-diversity and t-closeness. Wang et al.[21] proposed a location privacy preservation method based on k-anonymity and Voronoi maps, which ensures the privacy of location information while guaranteeing the security of the process and high-quality service.

As the performance of mobile terminals keeps promoting, their calculation and storage capability are also improved greatly. As a result, privacy protection scheme based on mobile terminals becomes feasible, which may serve to solve the problem of performance bottleneck and of the security's dependency on the trusted anonymity center of existing schemes. Li et al. [22] proposed the use of hidden Markov transfer matrix model to predict the user's motion trajectory, and will be used The forecast position of the next moment is used as the query content of the previous moment; Yang et al.[23]propose a k-anonymous location privacy protection scheme via dummies and Stackelberg game. The proposed scheme can effectively resist the single-point attack and inference attack while balancing the service quality and location privacy. Despite that these schemes avoid the dependency on the trusted anonymity center, they still focus only on user's privacy, while ignore the privacy of LBS server. If single user is able to infer the overall information of LBS server based on the partial information obtained, other users' private information is at the risk of disclosure and LBS server may be invalid. Therefore, the privacy of LBS server may as well be protected.

2.2. Architecture of location privacy protection system

Based on K-anonymity and OT protocol, this paper devised an architecture of location privacy protection without a third party, which is mainly composed of two entities: MT and LBS server, as is shown in Figure. 1. The functions are as follows:

MT: sending anonymization request to LBS server; generating & selecting dummy location nodes, sending

location query request to LBS server and receiving query result.
LBS: server: dealing with user's registration and query of interest points, encrypting the query results and returning them to MT.

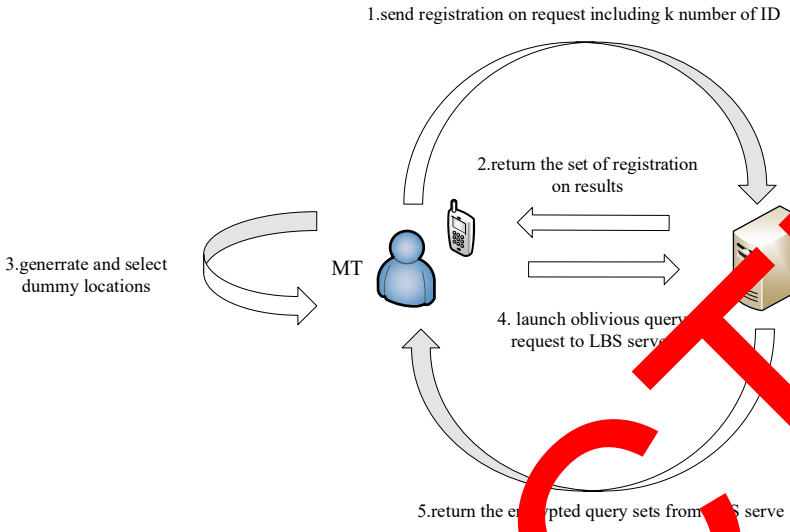


Figure 1. Architecture of location privacy protection system in this scheme

2.3. Position entropy

Given a cloaking region that contains number of candidate locations, the probability of each location becoming real location is marked as:

$$Pr(i) = \frac{Y_i}{\sum_{i=1} Y_i} \quad (1)$$

then its value of location entropy is:

$$H(\cdot) = \sum_{i=1}^{k} \Pr(i) \times \log \Pr(i) \quad (2)$$

Equations (1) and (2) can be used to obtain the location entropy of the candidate nodes, the higher the value is, the more secure the privacy protection can be. Obviously, when all are equivalent, the higher the entropy value of location nodes, the more secure the privacy protection can be.

3. K-anonymity Privacy Protection Scheme based on Terminal

3.1. System Initialization

In this phase, system parameters are generated as follows:
Step 1: Select two cyclic groups G_1 and G_2 with order q, in which G_1 is the addition cyclic group, G_2 the multiplication cyclic group, and q a big prime number. Let $e: G_1 \times G_1 \to G_2$ denote a bilinear pairing.
Step 2: Define three harsh function: H_1, H_2 and H_3, in which, $H_1: \{0,1\}^* \to G_1^*$, $H_2: G_2 \to \{0,1\}^n$; H_3 is the harsh function of SHA256; n denotes a integer, while $\{0,1\}^*$ is the binary string at any length.
Step 3: LBS server selects a random number $s \in Z_q^*$, let s be the system's private key, calculates its public key $PK = sP$, in which P is the generator of G_1.
Step 4: LBS server stores the system's private key s, publicizes the public parameter: $\{G_1, G_2, e, n, q, P, PK, H_1, H_2\}$.

3.2. User registration

In this phase, MT sends k number of ID (including itself) to LBS server as registration query message, then LBS server generates pseudonyms and corresponding public/private key pairs for these ID and returns them to users. The specific steps are as follows:
Step 1: MT sends k number of identity information $\{ID_1, ID_2, \ldots, ID_k\}$ to LBS server as registration query, in which user's identity ID_U is located at the u-th position in

the line of requested messages, $u \in \{1,2,\mathrm{L}\ ,k\}$, and u is selected by the user.

Step 2: LBS server employs a pseudorandom generator to generate salt figure ID_{salty}, and calculates separately the corresponding pseudonyms $PID_i = H_3(ID_i + ID_{salty})$, the corresponding public key $U^i_{PK} = H_1(ID_i)$, and the private key $U^i_{SK} = sU^i_{PK}$, then generates message $\{(PID_1, U^1_{PK}, U^1_{SK}), (PID_2, U^2_{PK}, U^2_{SK}), \mathrm{L}\ , (PID_k, U^k_{PK}, U^k_{SK})\}$, and returns it to user via safe channel.

Step 3: After receiving the message, the user calculates $U_{PK} = H_1(ID_U)$, and decides whether the equation $e(U_{SK}, P) \stackrel{?}{=} e(U_{PK}, PK)$ is valid or not. If valid, the user obtains correct dummy ID, public key and private key; if not, user gets invalid message, and returns to Step 1.

3.3. Generation and selection of dummy locations

The user B generates $2k$ number of dummy locations via MT, from which $k-1$ number of more favorable locations are selected. The specific procedure is as follows:

Step 1: Centering around L_B ---the location of the user B, MT generates a dummy location L_i by employing algorithm of uniformly distributed random points in rectangular region. Suppose the rectangular region is $[a,b]\times[e,f]$, generates independently the uniform random number α_i and β_j within $(0,1)$, then calculates $x_i = (b-a)\beta_i + a$ and $y_i = (f-e)\alpha_i + e$, so as to obtain respectively the random position $L_i = (x_i, y_i)$ are uniformly generated in $[a,b]$ and $[e,f]$. Then judge the correctness of the location based on mapping background information. If the location is a mountain or river, discard it and generate a new location; otherwise, calculates Euclidean distance $dis(L_B, L_i) = \sqrt{(x_B - x_i)^2 + (y_B - y_i)^2}$ between this location and location $L_B = (x_B, y_B)$.

Step 2: Step 1 judges whether equation $R_{min} < dis(L_B, L_i) < R_{max}$ is valid or not, where R_{min} and R_{max} respectively denote for the shortest and longest distance from center to newly-generated dummy location. If valid, let c_i where is a member of the false position set C, and add it into location set $C = \{c_1, c_2, c_3, \mathrm{L}\ , c_{i-1}\}$, that is $C = \{c_1, c_2, c_3, \mathrm{L}\ , c_{i-1}\} \cup c_i$; if not, return to Step 1.

Step 3: If $i < 2k$, then $i=i+1$, and return to Step 1; if $i=2k$, go to next step.

Step 4: Based on formula (1), calculate $Pr(i)$ ---the probability of each dummy location becoming real location, then select from $2k$ number of false locations $k-1$ number of false locations with higher $pr(i)$ --- $\{L_1, L_2, ...L_{K-1}\}$, to form more favorable dummy location set C_{End}.

Step 5: MT allocates fake identities to k number of location nodes (including MT itself).

3.4. Location service request

In this phase, MT sends location service request to LBS server, and LBS server answers the service request. The steps are as follows:

Step 1: LBS server randomly selects $d_1, d_2, \mathrm{L}\ , d_n \in Z^*_q$, in which $n \geq k$ calculates separately $P = d_1 PK$, $P_2 = d_2 PK, \mathrm{L}\ , P_n = d_n PK \in Z_q$ and releases them as basic points of location.

Step 2: The user B randomly selects $a_1, a_2, \mathrm{L}\ , a_k \in Z^*_q$, and calculates $v_i = a_i P_i$ in which $i = 1,2,\mathrm{L}\ ,\ k$.

Step 3: Gather the generated fake identities, location node and dummy query information to form a query set: $Msg = \{(PID_1, L_1, Q_1, v_1), (PID_2, L_2, Q_2, v_2), \mathrm{L}\ , (PID_u, L_u, Q_u, v_u), \mathrm{L}\ , (PID_k, L_k, Q_k, v_k)\}$ then send it to LBS server, in which L_u is the location of real node, and Q_u is the real query.

Step 4 After receiving location service request, LBS server obtains k number of query results $\{m_1, m_2, \mathrm{L}\ , m_u, \mathrm{L}\ , m_k\}$, then selects random number $r \in Z^*_q$, calculate $Y_0 = rPK$, $Y_i = rv_i$, and $c_i = m_i \oplus H_2(e(P_i + sPK, U^B_{PK})^r)$, finally sends $\{Y_0, (Y_1, Y_2, \mathrm{L}\ , Y_u, \mathrm{L}\ , Y_k), (c_1, c_2, \mathrm{L}\ , c_u, \mathrm{L}\ , c_k)\}$ to the user B. U^B_{PK} and U^B_{SK} respectively denotes the public and private key of the user B.

Step 5: After receiving the message, the user B judges whether the equation $e(v_u, Y_0) = e(Y_u, PK)$ is valid. If valid, the user B calculates the multiplication inverse $a^{-1}_u \in Z^*_q$ of a_u, and computes as well as decrypts the secret key $V_u = a^{-1}_u Y_u$, so as to obtain query results by figuring out $m_u = c_u \oplus H_2(e(V_u, U^B_{PK})e(Y_0, U^B_{SK}))$ (for multiple query requests, calculate multiple query results); if not valid, the user B discards this query result, returns to Step 1 and re-starts service request.

4. Scheme Analyses

4.1 Correctness Analysis

The correctness of this scheme can be proved if the user can succeed in obtaining the query result m_u.

When receiving the responded result $\{Y_0, (Y_1, Y_2, \mathrm{L}\ , Y_k), (c_1, c_2, \mathrm{L}\ , c_k)\}$ from LBS server, the

user firstly calculates the multiplication inverse $a_u^{-1} \in Z_q^*$ of a_u, calculates and decrypts the secret key $V_u = a_u^{-1} Y_u = a_u^{-1} r v_u = a_u^{-1} r a_u P_u = rP_u$, then verifies the correctness of $m_u = c_u \oplus H_2(e(V_u, U_{PK}^B) e(Y_0, U_{SK}^B))$.

Since
$$\begin{aligned} & c_u \oplus H_2(e(V_u, U_{PK}^B) e(Y_0, U_{SK}^B)) \\ &= c_u \oplus H_2(e(rP_u, U_{PK}^B) e(rPK, U_{SK}^B)) \\ &= c_u \oplus H_2(e(rP_u, U_{PK}^B) e(rPK, sU_{PK}^B)) \\ &= c_u \oplus H_2(e(P_u, U_{PK}^B)^r e(PK, sU_{PK}^B)^r) \\ &= c_u \oplus H_2(e(P_u, U_{PK}^B)^r e(sPK, U_{PK}^B)^r) \\ &= c_u \oplus H_2(e(P_u + sPK, U_{PK}^B)^r) \end{aligned}$$

And $c_u = m_u \oplus H_2(e(P_u + sPK, U_{PK}^B)^r)$,
$c_u \oplus H_2(e(V_u, U_{PK}^B) e(Y_0, U_{SK}^B)) = m_u$.

4.2. Security Analysis

4.2.1 Anonymity

Definition 1: Anonymity game

Step 1: The attacker A launches query to obtain the system parameters: $\{G_1, G_2, n, q, PK, H_1, H_2\}$, and LBS server publicizes the selected basic points $P_1, P_2, \L, P_n$ and $Y_0, Y_1, Y_2, \L, Y_k$;

Step 2: The attacker A selects k number of total different messages $m_1, m_2, \L, m_k$ as the candidate query results of the user B;

Step 3: The user B selects the random $u \in \{1, \L, k\}$, then send $\{m_1, m_2, \L, m_u, \L, m_k\}$ to LBS server, while u is not open to the attacker A;

Step 4: LBS server encrypts $\{m_1, m_2, \L, m_k\}$ and returns the corresponding result $\{c_1, c_2, \L, c_k\}$ to user B;

Step 5: If the encrypted results $\{c_1, \L, c_u, \L, c_k\}$ and $\{m_1, m_2, \L, m_k\}$ received by user B correspond with each other, then sends them to the attacker A in random order, otherwise, cease the game;

Step 6: If the attacker A decrypts ciphertext results and outputs $m_u' = m_u$, then he wins the game.

The advantage of the attacker A winning this game is: $Adv(A) = |Pr[A]|$, in which $Pr[A]$ represents the probability of attacker A outputting the message $m_u' = m_u$, that is, the probability of attacker obtaining user's real query results.

Theorem 1: In this scheme, assume attacker A is able to win the anonymity game with negligible probability, then this scheme satisfies the requirement for anonymity.

Proof: Suppose the attacker A obtains k number of ciphertext results $\{c_1, c_2, \L, c_u, \L, c_k\}$, in which $c_u = m_u \oplus H_2(e(P_u + sPK, U_{PK}^B)^r)$, while $r \in Z_q^*$ is the random number generated temporarily by LBS server, and U_{PK}^B is the public key of the user B. Suppose the attacker A attempts to obtain user's privacy information m_u by decrypting the ciphertext, then he has to solve $H_2(e(P_u + sPK, U_{PK}^B)^r)$, which means A must obtain P_u ---the basic point selected by the user, s ---the private key of LBS server, and r ---the random number generated temporarily. Since $P_u = d_u PK$, and d_u is selected randomly by LBS server and is unknown to the attacker A, so A cannot guess the basic result P_u that is corresponding with the real location of the user B; further, the attacker A he attempts to obtain the secret message (r, s) via $PK = sP$ and $Y_0 = rPK$, then solving the problem must be solving Elliptic Curve Discrete Logarithm Problem (ECDLP), which is infeasible in calculation. Consequently, the probability of attacker A obtaining the privacy message m_u of the user B is $Adv(A) = |Pr[A]|$, which is negligible. Therefore, the scheme meets the requirement of anonymity.

4.2.2. Resistance to replay attack

Definition 2: he attacker A re-sends the user's request message for registration and location service which have been processed by LBS server, so as to obtain the same results as the user B. The information known to the attacker A includes: $\{G_1, G_2, e, n, q, P, PK, H_1, H_2\}$ ---system's public parameters, $\{ID_1, ID_2, \L, ID_k\}$ --- the registration request message of the user B, $Msg = \{(PID_1, L_1, Q_1, v_1), (PID_2, L_2, Q_2, v_2), \L, (PID_k, L_k, Q_k, v_k)\}$ ---the request message for location service, and $P_1, P_2, \L, P_n$ --- the selected basic point that LBS server publicizes.

Theorem 2: if the attacker A obtains the same registration results and location service request results as the user B with negligible probability, then this scheme is able to resist replay attack.

Proof: it is known that the result of the user's registration request is $\{(PID_1, U_{PK}^1, U_{SK}^1), (PID_2, U_{PK}^2, U_{SK}^2), \L, (PID_k, U_{PK}^k, U_{SK}^k)\}$, in which the pseudonym is $H_3(ID_i + ID_{salty})$. Since the salt value ID_{Salty} is added in generating the pseudonym, ID_{Salty} is the one-time random number generated by LBS server by means of pseudo-random number generator based on encryption. Therefore, even if the attacker obtains the same identity information to resend registration request, he still cannot acquire the identical pseudonym, which means he cannot obtain the same results from registration request.

It is known that the results of the user's registration request is $\{Y_0, (Y_1, Y_2, \L, Y_k), (c_1, c_2, \L, c_k)\}$, in which $Y_0 = rPK$, $Y_i = rv_i$, $v_i = a_i P_i$ $(i = 1, 2, \L, k)$. Since r and a_i are both the one-time random number generated

randomly by LBS server, so the attacker still cannot obtain the same Y_0 and $Y_i(i=1,2,\mathrm{L},k)$ as the user, even if he obtains the location request message to replay; while $c_i = m_i \oplus H_2(e(P_i + sPK, U_{PK}^B)^r)$, in which r is the one-time random number, so the attacker cannot obtain the same ciphtertext $(i=1,2,\mathrm{L},k)$ from the location service request. Even if the attacker obtains the ciphtertext c_i, he still has to decrypt it.

Through the analysis above, if the attacker A attempts to obtain the same registration results and location service results as the user B via replay, the probability is negligible, so this scheme is able resist replay attack.

4.2.3. Non-forgeability

Let H_1 be the random oracle model from $\{0,1\}^*$ to G_1^*, then under the corresponding Random Oracle Model (ROM), devise a non-forgeability game based on chosen plaintext attack. The two sides in the game are the challenger C and the attacker A, and the model operates as follows:

(1) Initialization: the challenger C generates open systematic parameter $\{G_1, G_2, e, n, q, P, PK, H_1, H_2\}$ and confidential master key s.

(2) Training stage: the attacker A sends identity information ID_i to the challenger C, and requests an answer from oracle model H_1, whereas the challenger C performs key generation algorithm to generate and return the corresponding public/private key pairs to the attacker A. This stage can be repeated polynomial bounded times.

(3) Challenge stage: the attacker A randomly specify a user's identity ID_u as attack target---ID_u appear in any query process during the training stage, and this unknown corresponding public/private key pairs of the user are $U_{PK} = H_1(ID_u)$ and $U_{SK} = sH_1(ID_u)$ respectively. Under the condition that the user's public key is known, the attacker forges the user's private key U'_{SK}, attempting to make the equation $e(U_{PK}, P) \stackrel{?}{=} e(U_{SK}, PK)$ valid. If valid, then the attacker wins the game if not, this encryption algorithm satisfies non-forgeability.

Theorem 1: In the ROM, this scheme, if the attacker A cannot find the solution LBS server's private key s and provisional session's private key r in polynomial time, then this scheme satisfies non-forgeability.

Proof: During initialization, the challenger C discloses to the attacker A the parameter $\{G_1, G_2, n, q, P, PK, H_1, H_2\}$ and the communications messages between the user and LBS server, in which $PK = sP$, while s is the random number generated by the system and in unknown to the attacker A.

During the stage of user's registration, the attacker launches adaptability query to ROM H_1, inquires and obtains the corresponding harsh values. The process is as follows:

The attacker A asks the challenger C for the harsh value $H_1(ID_i)$ of the identity information ID_i, then the challenger C checks whether it exists in the request-reply list:

(1) if exist, return the corresponding reply to the attacker A;

(2) if not, generate randomly $\tau_i \in Z_q^*$, calculate $H_1(ID_i)$, then send $\{\tau_i, H_1(ID_i)\}$ to the attacker A, and store this request-reply in its list.

In the process of inquiry, the attacker A cannot obtain the system private key s, invariably cannot obtain the equation $e(U_{SK}, P) \stackrel{?}{=} e(U_0, PK)$ valid. If the attacker attempts to infer the system private key s from the open system parameter $\{P, PK\}$ and $PK = sP$, he will be confronted with solution of the elliptic curve discrete logarithm problem (ECDLP). Similarly, during the stage of location service request, the provisional session key r chosen by the system is unknown to the attacker A, thus the attacker A cannot forge correct Y'_u to make the equation $e(Y_u, Y_0) \stackrel{?}{=} e(Y'_u, P)$ valid. If the attacker A attempts to infer r from $Y_0 = rP$, then he will be confronted with solution of elliptic curve discrete logarithm problem (ECDLP).

5. Simulation Experiment

5.1 Efficiency

5.1.1 Communications costs

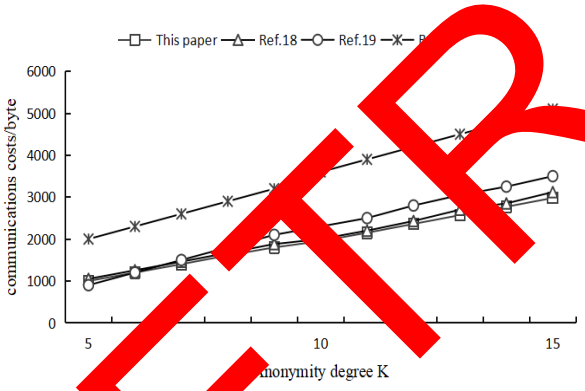
In an intact location service request, the communications costs mainly come in the stages of user registration and location service request, because those are the only places where information is exchanged, and the major communications data package includes: message of user's registration request, results of user's registration, message of location service request and the set of encryption results. The amount of communications costs depends on the size of the data package and anonymity degree K. The selection range of the parameter K in this experiment is [5, 15], as is shown in Fig. 2, and there is a linear relationship between communications costs and anonymity degree K in this scheme and other schemes. Since the scheme in reference[20] is based on the third-party anonymity center, and its communications must go through the center, so its communications costs increase with increasing anonymity degree K, at a faster speed than other two schemes. reference [18] in the scheme to construct the dummy anonymous set, need to verify the user's historical query probability, at the same time, discrete selection, so its communication time is slightly higher than the present scheme. reference [19] only submits a perturbation location

for querying, so its communication overhead is slightly lower than this scheme.

As a result, the results of this experiment demonstrate that this scheme can reduce the communications costs in the existing K –anonymity technology, thus has a certain advantage.

5.1.2. execution time

In this simulation experiment, program efficiency is measured by the time required to execute the algorithms proposed in each program. And the time cost of executing the algorithms in this scheme and the comparison scheme occurs mainly in the virtual location generation and selection phase. Since the magnitude of anonymity determines the number of virtual locations to be generated as well as the optimal virtual location, the execution time varies with the anonymity K. And the selection range of parameter K is assumed to be [10, 80] in this experiment. As shown in Fig. 3, when the value of anonymity K is taken low, the execution time of this scheme is close to that of reference [18] [20] and slightly higher than that of reference [19] but with the increase of K, the execution time of reference [19] is gradually higher than that of this paper. Reference [18] constructs an anonymous set while requiring discrete selection of locations, so its running time is gradually higher than this method. The execution time in reference [20] increases in a stable linear fashion a[s] increases. However, the algorithm in reference [20] is m[ore] complex and less efficient to execute than this scheme.



Fig[ure] 2. Th[e] relati[on]ship between degree of a[no]nymity K [and] communications costs

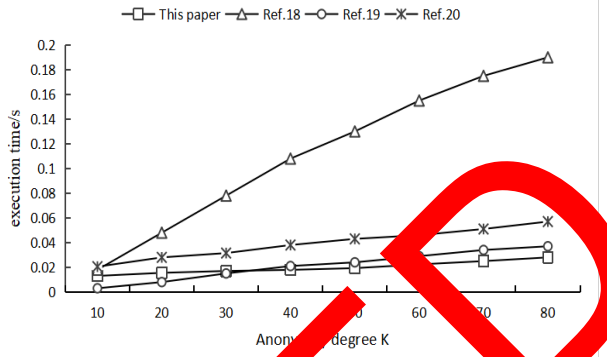


Figure 3. The [re]lations[hip] [b]etween a[no]nymity de[gr]ee K and ex[ecu]tion ti[me]

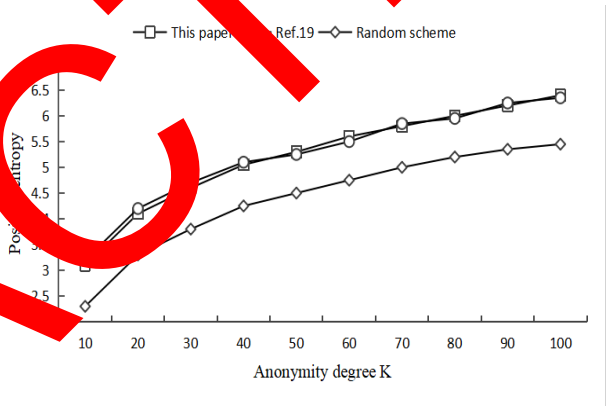


Figure 4. The relationship between anonymity degree K and position entropy

5.2. Privacy level

The location privacy level is usually measured by location entropy, measuring the quality of the chosen candidate nodes in the stage of dummy location generation and selection. The higher the entropy value is, the better the privacy is protected. With the occurrence probability of all location nodes (including the real location) being equivalent, the location entropy reaches its maximum level ideally. The higher the anonymity degree K is, the more confusing the user's real location will be; but if K is excessively high, the communications costs and efficiency will be affected, so the value of anonymity degree K is assumed as [10, 100]. As is shown in Fig. 4, the privacy level in three schemes all invariably increases with increasing anonymity degree K . However, when the value of K reaches a certain level, the increasing speed tends to be mild. That is because the obfuscation ability tends to be saturated as the dummy locations in the cloaking region become excessively dense, so they have little effect in

protecting the privacy no matter how many dummy locations are added. Since the random scheme doesn't takes into account the rationality of map information and dummy locations, so it has the most unfavorable effect in privacy protection. Although reference [19] fully considered the query probability of interest points in the process of generating anonymous sets, and selected interest points with similar probability of querying the user's location to constitute the anonymous set, its privacy effect is still inferior to that of the present scheme, which is lower than the present scheme by 0.2 %.

6. Conclusions

This paper proposes a K-anonymity location privacy protection scheme based on mobile terminal, which adopts secure and efficient m-out-of-n oblivious transfer protocol, thus avoids the dependency on the trusted anonymity center in existing schemes, improves the execution efficiency, and reduces the communications costs. Moreover, this scheme selects randomly $2k$ number of rational dummy locations from the cloaking region, then continues to select $k-1$ number of more favorable dummy locations according to locations entropy, thus improves the privacy level. The security analyses demonstrate that this scheme satisfies such security properties as anonymity, resistance to replay attack and non-forgeability. And the simulation experiment is also conducted to verify the communications costs, execution efficiency and privacy level, and the results show that the proposed scheme is superior to other schemes. Therefore, the scheme is of important theoretical significance and applicable value in security research related to location privacy protection. For future work, it will consider applying deep learning models to location privacy to further enhance security.

Acknowledgements.
This work was supported by the National Natural Science Foundation of China (61872126, 62273290); the Science and Technology Research Program of Henan Province (212102210092, 222102210078)

References

[1]. Yadav V K, Andola N, Verma S, et al. Anonymous and Linkable Location-Based Services[J]. IEEE Transactions on Vehicular Technology, 2022, 71(9): 9397-9409.
[2]. Song C, Zhang Y, Gu X, et al. A trajectory substitution privacy protection scheme in location-based services[J]. KSII Transactions on Internet and Information Systems (TIIS), 2019, 13(9): 4771-4787.
[3]. Yang G, He Y, Xiao K, et al. Privacy-Preserving Query Scheme (PPQS) for Location-Based Services in Outsourced Cloud[J]. Security and Communication Networks, 2022, Article ID 9360899.
[4]. Jin Y P. The investigation and analysis report on personal privacy data leakage in the era of big data[J]. Journal of Tsinghua University (Philosophy and Social Sciences), 2021, 36 (1): 191-201, 206.
[5]. Shen Z C, Zhang Q L, Zhang C F, et al. Location privacy attack based on deep learning [J]. Journal of Computer Research and Development, 2022, 59(2):390-402.
[6]. Zhang Y D, Smart detection on abnormal breasts in di mammography based on contrast-limited adap histogram equalization and chaotic adaptive real-cod biogeography-based optimization, Simulation, 2016, 92(9) 873-885
[7]. Zhang Y D, Feature Extraction of Brain by tionary Wavelet Transform and its Appli ions, Jo nal of Biological Systems, 2010, 18(s1): -132
[8]. Gedik B, Liu L. Protecting ation P Personalized k-Anonymi hit ure and Algorithms[J]. IEEE Tra ctions on le Computer, 2008, 7(1): 1-18.
[9]. Wang K, Zhao W, C J, l. A K-anonymo lustering algorithm based the a tic hierarchy process[J]. Journal of Visual Comm ication and Image Representati , 2019, 59: 76-83.
[10]. Memon I hen L, Arain Q A, et al. udonym changing strate with multi mix zones for trajectory privacy prote n in ro etworks[J]. International Journal of Commu tic ystems, 201 31(1): e3437.
[11]. hou B, Pe The k-anon ty and l-diversity approaches fo rivacy servati in social networks against neig orhood a]. Knowledge and Information Sys ms, 2011, 28(47-77.
[12]. I mavi R. Optimization-based k-anonymity lgorithms. Computers & Security 2020, 93: 101753.
[Zhang L, Jin C, Huang H P, et al. A Trajectory Privacy reserving Scheme in the CANNQ Service for IoT [J]. ors, 2019, 19(9).
[14]. Zhang S, Mao X, Khoo K K, et al. A trajectory privacy-preserving scheme based on a dual-K mechanism for continuous location-based services[J]. Information Sciences, 2020, 527: 406-419.
[15]. Khan, M A, Ullah I, Alkhalifah A, et al. A provable and privacy-preserving authentication scheme for UAV-enabled intelligent transportation systems[J]. IEEE Transactions on Industrial Informatics, 2021 18(5): 4316-3425.
[16]. Kuang L, Wang Y, Zheng X, et al. Using location semantics to realize personalized road network location privacy protection[J]. EURASIP Journal on Wireless Communications and Networking, 2020, 2020(1): 1-16.
[17]. Tu Z, Zhao K, Xu F L, et al. Protecting tr y from semantic attack considering k-anonymity versity, d t-closeness[J]. IEEE Transactions on N ork and Serv Management, 2019, 16(1): 264-278.
[18]. Yang Y, Hu X H, Du Y W. k- ymous dummy selection algorithm based on h rical qu robability J] . Computer Engineeri , 2022, 48(2): 7-155
[19]. Yan G H, Liu T, Zha X J, et al Service rity location k anonymity ivacy p ction scheme against background kno dge n ttack[J]. Jo al of Xi'an Jiaotong Univ ty, 2020, (1): 8-18
[20]. Zhang S B Q, Wang G J. ctor ivacy protection method sed location obfusc]. Chinese Journal on Co municati 2018, 39(07): 81-91
[21]. Wang B , Guo Y , et al. k-anonymity based location protection me d for location-based services in Internet of Thing[J].Concu ncy and Computation: Practice and Experience, 2021.DOI:10.1002/cpe.6760.
[22]. Li H W, Di S, Meng J J, et al. Spatio-temporal aware privacy-prese ng scheme in LBS[J] . Journal on Communicati s, 2018, 39(5): 134-142.
[23] ang D D e B P, Zhang W Y, Zhou H Y, Qian X B, "K A k-Anonymous Location Privacy Protection Scheme via Dummies and Stackelberg Game", Security and mmunication Networks, vol. 2021, Article ID 9635411, 15 pages, 2021. https://doi.org/10.1155/2021/9635411
[24]. Wang S H, Grad-CAM: understanding AI models, CMC-Computers, Materials & Continua, 2023, 76(2): 1321-1324
[25]. Xing L, Zhang D X, Wu H H, et al. Distributed K-Anonymous Location Privacy Protection Algorithm Based on Interest Points and User Social Behavior[J]. ELECTRONICS,2023,12(11).