Mobile Networks and Applications (2022) 27:1162–1173
https://doi.org/10.1007/s11036-022-01950-6

Intelligent Privacy Protection of End User in Long Distance Education

Yating Li[1,2] · **Jiawen Zhu**[1,3] · **Weina Fu**[1,2]

Accepted: 9 February 2022 / Published online: 18 March 2022

Abstract
Long distance education is an important part during the COVID-19 age. An intelligent privacy protection with higher effect for the end users is an urgent problem in long distance education. In view of the risk of privacy disclosure of location, social network and trajectory of end users in the education system, this paper deletes the location information in the location set to protect the privacy of end user by providing the anonymous set to location. Firstly, this paper divides the privacy level of social networks by weighted sensitivity, and collects the anonymous set in social networks according to the level; Secondly, after the best anonymous set is generated by taking the data utility loss function as the standard, it was split to get an anonymous graph to hide the social network information; Finally, the trajectory anonymous set is constructed to hide the user trajectory with the l-difference privacy protection algorithm. Experiments show that the algorithm presented in this paper is superior to other algorithms no matter how many anonymous numbers there are, and the gap between relative anonymity levels is as large as 5.1 and 6.7. In addition, when the privacy protection intensity is 8, the trajectory loss rate presented in this paper tends to be stable, ranging from 0.005 to 0.007, all of which are less than 0.01. Meanwhile, its clustering effect is good. Therefore, the proportion of insecure anonymous sets in the algorithm in this paper is small, the trajectory privacy protection effect is good, and the location, social network and trajectory privacy of distance education end users are effectively protected.

Keywords Intelligent protection · Long distance education · End user · Intelligent education system · Social network · Anonymous set

1 Introduction

Intelligent distance education is a new education mode, which fully combines the advantages of computer technology and communication technology to improve the quality and effect of intelligent distance education [1]. Traditional teaching has the limitations of resources such as teachers and environmental conditions, and the educational effect is general. Intelligent distance education breaks these limitations. Teachers and students from all over the country can interact and discuss through the intelligent distance education system [2], improve the quality of education and improve students' learning results [3].Intelligent distance education system can provide students with sufficient learning time and mobilize their learning enthusiasm. The advantages of intelligent distance education system are not limited by time and space, the realization of personalized and autonomous learning [4], and high sharing rate of educational resources.

In the context of big data, although the intelligent distance education system can improve the education effect to the greatest extent, its disadvantage is the security problems [5]. End users can learn anytime and anywhere through smart devices such as mobile phones and tablets in intelligent distance education systems [6].These terminal devices provide convenience for distance education, but they are also prone to the risk of disclosure of private information. Therefore, in order to ensure the security of end-user privacy information, it is necessary to study the end-user privacy protection

Yating Li and Jiawen Zhu These authors contribute equally.

✉ Weina Fu
fuwn@hunnu.edu.cn

[1] Hunan Provincial Key Laboratory of Intelligent Computing and Language Information Processing, Hunan Normal University, Changsha 410081, China

[2] College of Information Science and Engineering, Hunan Normal University, Changsha 410081, China

[3] College of Educational Science, Hunan Normal University, Changsha 410081, China

algorithm [7] to ensure that user privacy information will not be disclosed.

Luo et al. considered the distributed k-anonymous privacy protection method that does not take into account the credibility of the participants, so they proposed a trust-based location privacy protection algorithm. The trust management method based on Dirichlet distribution ensures that both requester and collaborators will only cooperate with vehicles they trust. Record the credibility of vehicles in public available blocks through the data structure of the blockchain, determine anonymous stealth areas, complete location privacy protection and improve the effect of location privacy protection. However, this algorithm lacks personalization and differentiation, and it will lose more data utility while protecting location privacy [8]. Wang et al. proposed a probability based source location privacy protection algorithm for wireless sensor networks (WSN) in view of the location privacy problem that is a research hotspot in security field, they estimates the state of the source through hidden Markov model, using virtual nodes and pseudo-sources to simulate the behavior of the source, and to realize the diversification of routing paths. In addition, two transmission modes are designed to transmit real data packets to realize location privacy protection. This algorithm shortens the time required by privacy protection without reducing energy consumption. However, this algorithm only considers the user's location privacy protection, and the privacy protection is not comprehensive [9]. Wu et al. first constructed a mathematical model, then solved the conditional entropy and mutual information, and encrypted the solution results through the homomorphic cryptosystem to complete the privacy protection of social networks and accelerate the efficiency of privacy protection. However, this algorithm does not consider the sensitive attributes of social network information, which is prone to excessive encryption of social network privacy and increase the rate of information loss [10].

Therefore, this paper studies the end user privacy protection algorithm of intelligent distance education system, designed location protection, social network privacy levels, anonymity maps and end user trajectories, comprehensive protection of end-user privacy, which can improve the effect of end-user privacy protection and reduce the information loss rate.

2 End user privacy protection algorithm

2.1 The location information protection algorithm based on K-anonymous context awareness

The location k-anonymity model of context awareness is used to protect the location privacy of end users in intelligent distance education system, which solves the problem of poor anonymity of location privacy information when there are few end users. Adding context awareness to the k-anonymity model can obtain the scenarios of end users, automatically control location privacy information according to different scenarios, and improve the effect of location privacy protection.

Context awareness is used to obtain the characteristics of the end user's current location and population density [11], which is used to represent the demand degree of the end user's location information privacy protection, and anonymously process the end user's location information according to the demand degree [12, 13]. The measure of end user's location context awareness is the location context and location offset δ, the distance between the end user's current location point and the query point is δ. The expression formula of end-user location privacy protection is as follows:

$$K = K_{\max} + K_{\max} \Big/ \left(\frac{K_{\max}}{K_{\min}} e^{-\alpha \frac{\lambda}{\delta}} - e^{-\alpha \frac{\lambda}{\delta}} \right) \tag{1}$$

where the decay rate is α and the route network density of subspace is λ.

The k–anonymity model based on context awareness is designed according to K, so that the original location set of end users is $L = (l_1, l_2, \cdots, l_m)$, in which the generalized form of l_i is $[u_i, TI_i, (x_{\min}, y_{\min}), (x_{\max}, y_{\max}), A_i]$, and the ID of end users is u_i, and there is only one; The generalization range of end-user location is $(x_{\min}, y_{\min}), (x_{\max}, y_{\max})$; The time interval is TI_i; The set of semantic identifiers is A_i.

When the model conceals the location information of anonymous subsets, it is affected by both spatial distance and time. In the time and spatial dimensions, the shorter the distance between interference location information and the location information that needs privacy protection, the better the effect of anonymous group on location privacy protection [14]. By measuring the similarity of location information in spatial and temporal dimensions, the effect of location privacy protection is improved. The similarity formula of location information in joint temporal and spatial dimensions is as follows:

$$\begin{aligned} S'''(a,b) = & \frac{w_1 * S'(T_a, T_b) + w_2 * S''(L_a, L_b)}{S'(T_a, T_b)} \\ & + \frac{w_1 * S'(T_a, T_b) + w_2 * S''(L_a, L_b)}{S''(L_a, L_b)} \end{aligned} \tag{2}$$

The location information are a and b; The weight of location information in time and spatial dimensions are w_1 and w_2; The similarity of time dimension location information is $S'(T_a, T_b) = 1 \Big/ \left\| Hash(T_a) - Hash(T_b) \right\|_2$, the hash values of time points a and b are $Hash(T_a)$ and $Hash(T_b)$. The similarity of spatial dimension position information is

$S''(L_a, L_b) = 1/\|L_a - L_b\|_2$, the location of a and b are L_a and L_b; The Euclidean distance between a and b is $\|L_a - L_b\|_2$

The principle of protecting the location privacy of end users of intelligent distance education system by using the location k-anonymity model based on context awareness is to search the other location set L' for L_i, and ensure that $u_i \neq u_j$ in each L_i. In the time and spatial dimensions, compare and analyze $S'''(a, b)$ of all location information in L', obtain the $K-1$ objects with closest similarity to L_i, generalize their location information in the time, spatial and points of interest dimensions, and establish an anonymous subset AS. All location information of each end user contains interest point identification, which is added to the hash table [15], and the hash table is used to measure whether the end user's current location is included in the anonymous location range. For each data point A_i in the anonymity subset, when the initial interest point identification of A_i and A_i is consistent, compare the anonymous space range and the new anonymous space range. The corresponding space range of the location is a small anonymous area with corresponding adjusted interest point attributes; when the initial interest point identification of A_i and A_i is inconsistent, the corresponding hash table needs to be adjusted. The specific operation steps of end user location privacy are as follows:

solve K according to the query records of the end user in the intelligent distance education system;
arrange the historical query records according to the time and spatial location to obtain the candidate L';
Solve position information $S'''(L_i, L_j)$ in L', arrange position information according to $S'''(L_i, L_j)$, select the first K information to establish $AS(L_i)$ of L_i, if A_i is consistent with the identifier of interest before anonymity, then compare and analyze the scope of anonymous space and new anonymous space scope, adjust $AS(L_i)$. Make the scope of the smaller space be the corresponding space scope of the location, and the interest identifier is changed. Solve L_j and other records $S'''(L_i, L_j)$, arrange position information according to $S'''(L_i, L_j)$ to obtain the candidate anonymous set L_a, the final anonymous set of location information is obtained by adjusting L' through L_a, and the location information in the anonymous set is deleted in L, thus the location privacy of the end user of the intelligent distance education system is realized.

2.2 Social network privacy ranking based on weighted sensitivity

In order to realize the protection of social network privacy of end-users in intelligent distance education system, it is necessary to define the level of social network privacy by using the anonymous model based on utility differentiation. The weighted sensitivity is used to complete the task.

First of all, complete the setting of relevant parameters as follows:

To make the end user attribute-the social network graph is $G = (V, E, B, H)$; the end user node set is V; the set of social relationship of end user is $E \subseteq V \times V$ wherein the type of all the relationship is the same, that is, they are not sensitive information [16]; the node set of sensitive attribute value is $B = \{B_1, B_2, \cdots, B_{n'}\}$, the mapping relationship between the end user node and $B_{n'}$ is H; the attribute tag Z of end user's social network contains $B_{n'}$ and the sensitivity of attribute value.

Then, by judging the sensitivity of $B_{n'}$, differentiated social network privacy protection is carried out according to the attribute values of different sensitive attribute categories, and reduce the information loss rate when processing social network information anonymously. The specific steps of social network privacy ranking based on weighted sensitivity are as follows:

when the end user sends transaction request information in the intelligent distance education system, according to the social network privacy protection requirements set by the end user, they can obtain the set of protected social network privacy information and the corresponding end user node V_o.
initialize the social network information of the end user, replace the unique identifier of the end user with serialization symbol, and establish G.
using the $n' \times (m+1)$ dimensional weighted sensitivity matrix $M|B$ to solve the correlation weighted sensitivity degree between identifiers and $B_{n'}$ in the end user's social network privacy information, in order to obtain the social network privacy level, and the social network privacy level is stored in Z and establish an anonymous set of social network. The expression of the weighted sensitivity calculation matrix is as following Eq. 3.

$$M|B = \left(t_{i'j'}|b_{i'}\right)_{n' \times (m+1)} = \left[\begin{array}{cccc|c} Q_1 & Q_2 & \cdots & Q_{n'} & B_1 \\ t_{11} & t_{12} & \cdots & t_{1n'} & b_1 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ t_{n'1} & t_{n'2} & \cdots & t_{n'm} & b_{n'} \end{array}\right] \tag{3}$$

where, the number of identifier Q attributes is m; the influence degree of the j' attribute of the i' tuple in the matrix $M|Y$ with respect to $B_{n'}$ is $t_{i'j'}$, and the sensitivity of $B_{i'}$ is $b_{i'}$; the influence degree of Q with respect to B_ρ is $F_\rho = \sum_{i'=1}^{|Q|} t_{\rho i'}$, where $|Q| = m$; the sensitivity degree of Q with respect to B_ρ is $b_{i'\rho} = \frac{t_{\rho i'}}{F_\rho}$.

Equation 3 can be used to describe the relationship between Q and $B_{n'}$, the importance of $b_{i'}$ and Q, and obtain the privacy level of social networks.

2.3 The establishment of anonymous graph

On the basis of obtaining the privacy level, take the data utility loss function as the standard to optimize the anonymous set to obtain the optimal anonymous set, and further split and deal with the optimal anonymous set to establish an anonymous graph to hide social network information; the steps are as follows:

obtain the optimal anonymous operation set. In the process parent nodes assign social relations to the child nodes, The minimum value T_C is obtained according to the data utility loss evaluation function $L'(G, G^*)$, which is regarded as the best standard for optimizing the anonymous set of social networks. To obtain the optimal anonymous operation set, the calculation formula of $L'(G, G^*)$ is as following Eq. 4.

$$L'(G,G^*) = \varepsilon_1 \cdot (T_C) + \varepsilon_2 \cdot (I) V_o s.t \begin{cases} T_C = \omega_1 \cdot \dfrac{P' - P}{P} + \omega_2 \cdot \dfrac{X' - X}{X} \\ \varepsilon_1 + \varepsilon_2 = 1 \\ \omega_1 + \omega_2 = 1 \\ \min L'(G, G^*) \end{cases} \tag{4}$$

Among them, the social network graph of end user after anonymously processing is G^*, the change of end user social network structure after anonymously processing is T_C, the information loss of end user social network is I, and the constants that measures T_C and I are ε_1 and ε_2. The average shortest path is P, the clustering coefficient is X, and the structural indices of evaluating G' are P' and X'. The weights of T_C and I are ω_1 and ω_2.

split the optimal anonymous set and establish an anonymous graph G' of the end user social network. In order to ensure the diversity and anonymity of the end user social network privacy, node splitting can be used to perform the steps of adding and deleting in the anonymous operation. The principle of node splitting is as follows:

$$D\left(V', u\right) \rightarrow \left\{V_1', u_1\right\} \cup \left\{V_2', u_2\right\} \tag{5}$$

Among them, the end user node to be split is V', that is, the parent node; the child nodes obtained by splitting V' are V_1' and V_2'. The 1-degree neighbor sub-graph of V' in G is u. The 1-degree neighbor subgraphs constructed by V_1' and V_2' are u_1 and u_2.

2.4 Design of anonymous set of end-user trajectory based on l-difference

The privacy protection algorithm based on trajectory l-difference in this paper realizes the track privacy protection of the end users of the intelligent distance education system. On the premise of conforming with the l-difference, make sure that the closely trajectories are anonymously connected. Make r_1 and r_2 in the time interval $\left[t_1', t_2'\right]$, the number of passing cells is θ_1 and θ_2. Set the threshold, that is, that is, the similarity ratio is l. When r_1 and r_2 are similar trajectories, then the trajectory correlation between r_1 and r_2 will exceed the $\left\lceil \min\left(\theta_1, \theta_2\right) \times l \right\rceil$. The l-difference of the end-user trajectory means that Under the similar ratio l, the random two trajectories r_1 and r_2 in the anonymous graph of the end user social network will not be similar trajectories, which ensures the diversity of the trajectories in ζ and avoids the disclosure of track privacy.

If G' has two trajectories r_1 and r_2, then r_1 and r_2 must not be similar, and the nodes τ_1 and τ_2 corresponding to r_1 and r_2 are connected to get an edge, so the set of end user trajectories can be changed into an undirected graph $h(\tau, \varpi)$. The two random trajectories in G' are not similar, which indicates that there is an edge connection between the corresponding nodes in g. Therefore, the clustering problem of k end user trajectories in search can be changed into a k-cluster problem which searches for the corresponding nodes wth clustering center trajectory in g. The specific steps of the privacy protection algorithm for trajectory l-difference are as follows:

Pr-process the track data set η of the end user in the intelligent distance education system, change η, and obtain several equivalent class data sets $\eta'\left(\eta' = \left\{\eta_1', \eta_2', \cdots, \eta_{m'}'\right\}\right)$, and the start and end time of all the end user trajectories in the equivalence class are the same.

Divide the two-dimensional space with all the trajectory sampling points of end user, obtain the cells [17], process $r_{a'}\left(a' \in \left[1, \left|\eta_{a'}'\right|\right]\right)$ in $\eta_{m'}'$ according to the order of cell segmentation, and obtain the cell list G' that $r_{a'}\left(a' \in \left[1, \left|\eta_{a'}'\right|\right]\right)$ passes.

On the premise of conforming with the l-difference, the $\eta'\left(\eta' = \left\{\eta_1', \eta_2', \cdots, \eta_{m'}'\right\}\right)$ in G' is processed by the clustering algorithm based on unified similarity measure, and the anonymous set of end-user trajectory privacy in intelligent distance education system is obtained. and delete the trajectories [18] in the anonymous set in η to realize the concealment of end-user track privacy.

Using the clustering algorithm based on unified similarity measure to cluster $\eta'\left(\eta' = \left\{\eta_1', \eta_2', \cdots, \eta_{m'}'\right\}\right)$, it is determined that the initial number of clustering is κ class [19]. Let the end user trajectory clustering itemset is

$B' = \{B'_1, B'_2, \cdots, B'_\kappa\}$, the similarity between η' and B' is $g(\eta', B')$. Then the objective function of clustering η' is shown in following Eq. 6.

$$W^* = \arg\max J(W, B') = \sum_{j''=1}^{\kappa} \sum_{i''=1}^{m'} \gamma_{i''j''} g\left(\eta'_{i''}, B'_{j''}\right) \tag{6}$$

where W is 0–1 matrix, $\gamma_{i''j''} \in \{0, 1\}$, $i'' = 1, 2, \cdots, m'$, $j'' = 1, 2, \cdots, \kappa$.

The track attribute of end user A' in η' is divided into numerical attribute x' and categorical attribute x'' [20]. The calculation formula of trajectory clustering similarity of mixed attributes is as follows:

$$g\left(\eta'_{i''}, B'_{j''}\right) = \frac{g\left(x'_{i''}, B'_{j''}\right) + t'_{B'} g\left(x''_{i''}, B'_{j''}\right)}{t'_f} \tag{7}$$

Among them, the similarity weight ratio of x' is $\frac{1}{t'_f}$ and the similarity weight ratio of x'' is $\frac{t'_{B'}}{t'_f}$.

The similarity between x' and x'' is solved by using Mahalanobis distance with Eq. 8.

$$\begin{cases} g'\left(x'_{i''}, B'_{j''}\right) = \dfrac{\exp\left(D'\left(x'_{i''}, B'_{j''}\right)\right)}{\sum_{\chi=1}^{\kappa} \exp\left(D'\left(x'_{i''}, B'_{\chi}\right)\right)} \\ g''\left(x''_{i''}, B'_{j''}\right) = \sum_{r'=1}^{t'_{C'}} \dfrac{\omega_{r'} \sigma_{A'_{r'}=x''_{i''r'}}\left(B'_{j''}\right)}{\sigma_{A'_{r'}\neq NULL}\left(B'_{j''}\right)} \end{cases} \tag{8}$$

Among them, $D'(\bullet)$ is the Mahalanobis distance, $\sigma_{A'_{r'}=x''_{i''r'}}\left(B'_{j''}\right)$ is the number of $x''_{i''r'}$ in B' and A', $\sigma_{A'_{r'}\neq NULL}\left(B'_{j''}\right)$ means the amount of A' in B' where the attributes number is not null, and the weight of A' is $\omega_{r'}$.

Using Eq. 8 into Eq. 7, $g\left(\eta'_{i''}, B'_{j''}\right)$ is calculated, and then $W^* = \left\{\gamma^*_{i''j''}\right\}$ is also calculated. When $\gamma^*_{i''j''} = 1$, $g\left(x'_{i''}, B'_{j''}\right) \geq g\left(x'_{i''}, B'_{r'}\right)$; when $\gamma^*_{i''j''} = 0$, $g\left(x'_{i''}, B'_{j''}\right) < g\left(x'_{i''}, B'_{r'}\right)$.

In this way, each x is assigned to the cluster item with the highest similarity, and the frequency of x'' and center of x' in each cluster item are continuously adjusted to obtain the end ser trajectory clustering results of the intelligent distance education system until $W^* = \left\{\gamma^*_{i''j''}\right\}$ is stable. Build an anonymous set of end user trajectory privacy to achieve trajectory privacy protection.

The similarity coefficient $\phi_{r'}$ is used to solve $\omega_{r'}$, and the specific steps are as follows:

Define the unknown measure matrix of $\omega_{r'}$ by Eq. 9.

$$\left(\mu_{m'\kappa}\right) = \begin{bmatrix} \mu_{11} & \cdots & \mu_{1\kappa} \\ \vdots & \cdots & \vdots \\ \mu_{m'1} & \cdots & \mu_{m'\kappa} \end{bmatrix} \tag{9}$$

Solve the $\phi_{r'}$ between the end user trajectory evaluation vector and the comprehensive measure evaluation vector by using Eq. 10.

$$\phi_{r'} = \frac{1}{m'} \sum_{i''=1}^{m'} \sum_{j''=1}^{\kappa} \mu_{i''r'j''} \mu_{i''j''} \tag{10}$$

Solve the $\omega_{r'}$ by using Eq. 11.

$$\omega_{r'} = \frac{\phi_{r'}}{\sum_{j''=1}^{\kappa} \phi_{r'}} \tag{11}$$

3 Experimental analysis

Taking a training institution as the research object, the number of end users of the training institution is 6×106. This algorithm is used to protect the privacy of the end users of the intelligent distance education system of the training institution, and the privacy protection effect of the algorithm is analyzed.

3.1 End user location privacy protection effect

The proposed algorithm needs to select a reasonable attenuation rate, which ensures that the privacy protection effect of the end user's location reaches the best state. In the case of different route network density and end user location offset ratio, the impact of attenuation rate on end-user privacy protection on privacy protection of end user is analyzed, and the results are shown in Fig. 1.

According to Fig. 1, in the case of different route network density and end-user location offset ratio, the protection intensity of location privacy increases with the attenuation rate, showing a trend of increasing at first and then declining. When the ratio is 0.2 to 0.4, the attenuation rate corresponding to the value of privacy protection of the highest location is 20%. When the ratio is 0.6, the corresponding attenuation rate of the privacy protection of the highest location is 15% to 20%. Comprehensive analysis shows that in order to ensure the maximum protection of location privacy, the attenuation rate is 20%, which can effectively improve the end user location privacy protection effect of this algorithm.

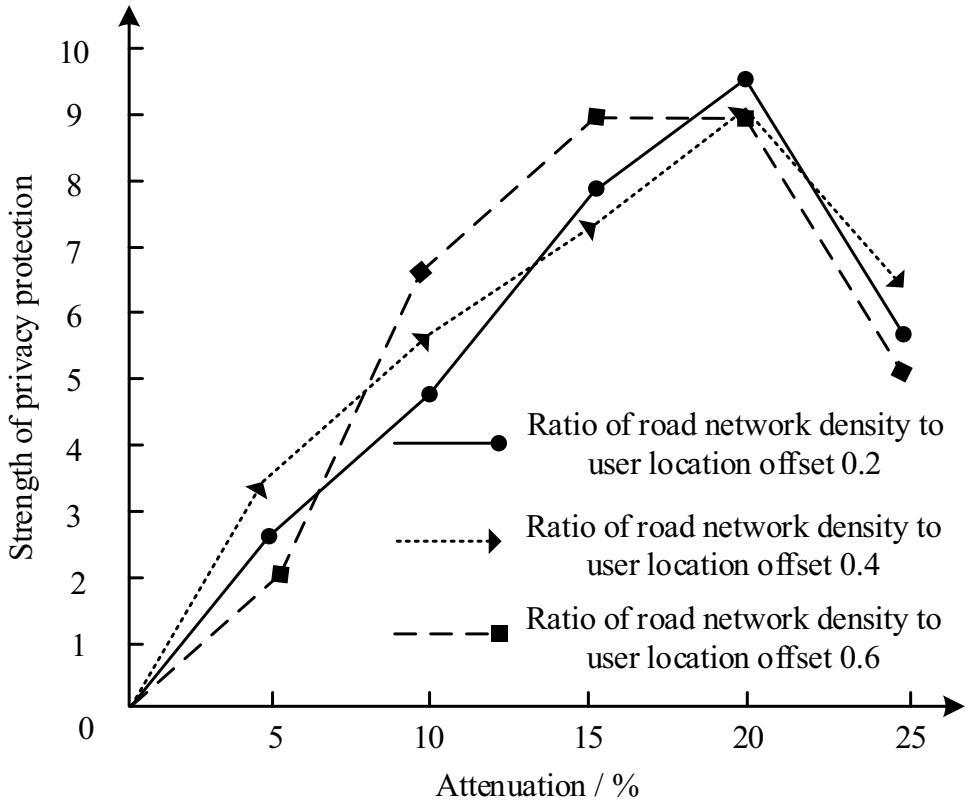

Fig. 1 Influence of attenuation rate on privacy protection of end users

Randomly select an end user in the training institution, the proposed algorithm anonymously deal with the location privacy of the end user in this paper, and analyze the protection effect of the algorithm on the location privacy of the end user. The processing results are shown in Fig. 2.

According to Fig. 2, the proposed algorithm can effectively anonymously deal with the location privacy information of the end user, ensure that all the location privacy information is controlled within the range of anonymous processing, and realize the concealment of the location privacy information of the end user. Experiments show that the proposed algorithm can effectively protect the location privacy of end users.

The comparison algorithm in this paper uses the trust-based location privacy protection algorithm in Reference 8, and the probability-based source location privacy protection algorithm in Reference 9. The Relative anonymous level(RAL) and information entropy are used to measure the effect of the three algorithms that protect end-user location privacy. RAL value is positively correlated with end-user location privacy protection effect. Through information entropy, the possibility of anonymity processing results being destroyed by an attacker is judged.The higher the value is, the lower the probability the attacker breaches the actual location of the end user is. When the number of anonymity is different, the RAL and information entropy test results of the algorithm to protect the location privacy of end users are shown in Figs. 3 and 4.

As can be seen from Fig. 3, RAL values of the three algorithms all increase with the increase of the number of anonymous people. Among them, the algorithm in this paper is higher than the other two algorithms when the number of anonymous is different. When the number of anonymous is 20, the relative anonymity level gap is the largest, reaching 5.1 and 6.7 respectively. Therefore, the algorithm in this paper has the best effect on location protection.

According to Fig. 4, the location entropy of the three algorithms all increase with the increment of the anonymity. When the number of anonymity is different, the location entropy of the algorithm in this paper is closest to the ideal value, while the location entropy of the other two algorithms differs from the ideal value, which shows that the algorithm in this paper has the best effect in protecting the location privacy of end users.

3.2 Privacy protection effect of end user social network

The average path length and clustering coefficient are used to measure the anonymous graph structure characteristics of the three algorithms after anonymous processing of end user social network privacy. the more similar the structure

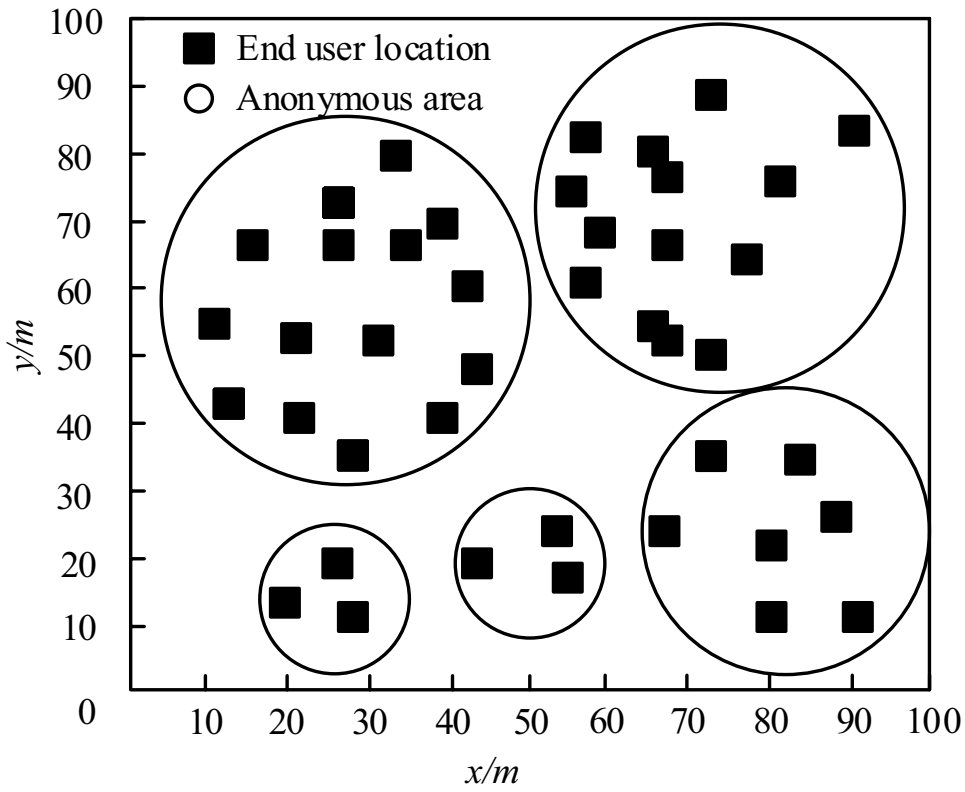

Fig. 2 Results of end user location privacy protection

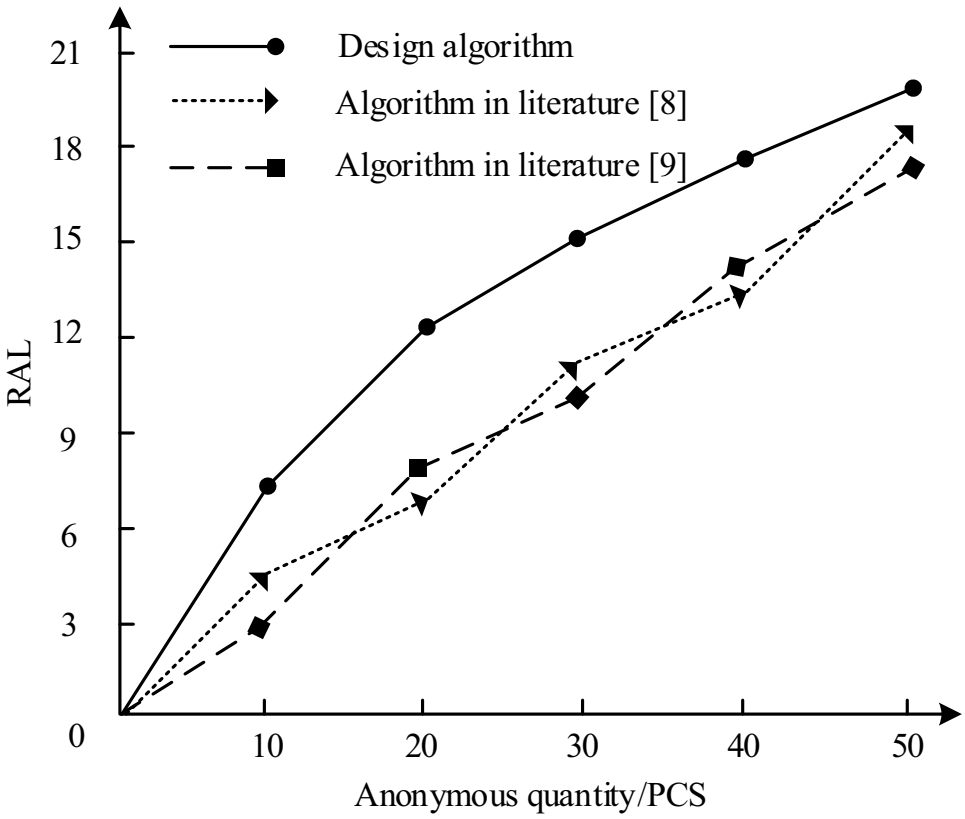

Fig. 3 RAL test results

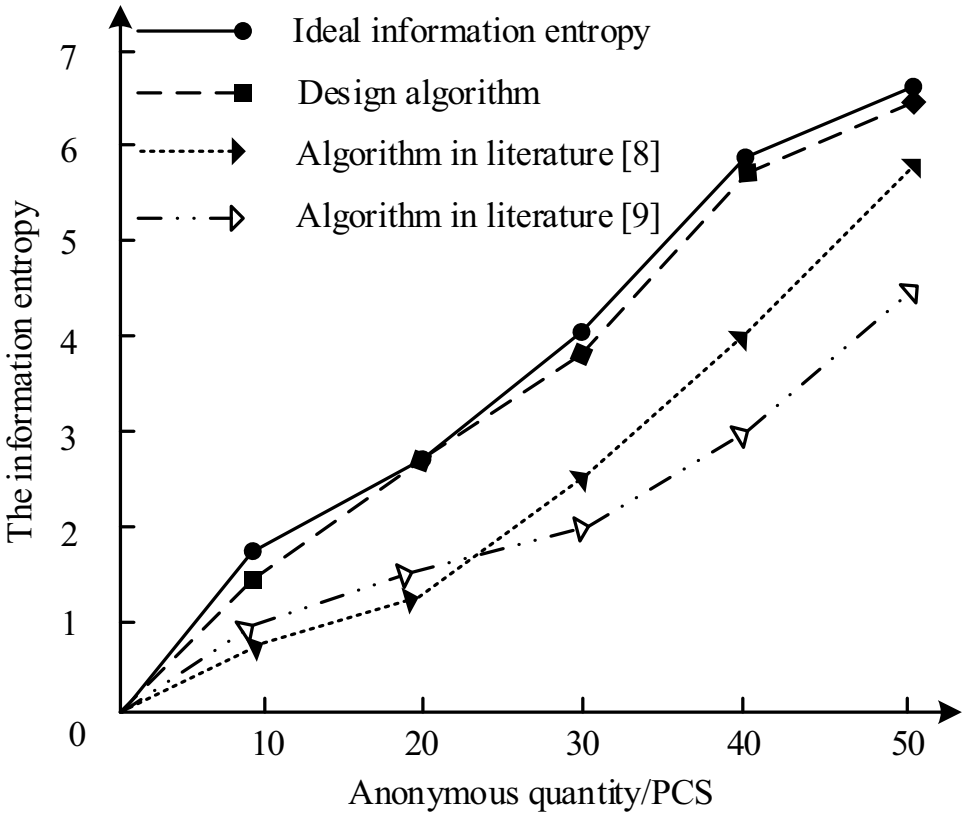

Fig. 4 Information entropy test results

characteristics of the anonymous graph are to the original graph, the less information loss is and the better the privacy protection effect of social network is. Test the average path length and clustering coefficient of social network anonymous graph with different number of anonymity of the three algorithms. In order to ensure the privacy protection effect of the end user social network, the test results are shown in Figs. 5 and 6.

According to Fig. 5, with the increase of the number of anonymity, the average path length of the anonymous graph constructed by the three algorithms shows a downward trend. When the number of anonymity is different, the average path length of the anonymous graph constructed by this algorithm, the difference between it and the average path length of the original graph is the smallest, and the average path length of the anonymous graph constructed by the other two algorithms is quite different from that of the original graph. The experimental results show that in the process of protecting the privacy of the end-user social network, the average path length of the anonymous graph constructed by this algorithm is only slightly lower than the average path length of the original graph, which is obviously better than the other two algorithms. It is shown that the structure of the anonymous graph of the end-user social network constructed by this algorithm is very similar to that of the original graph, which can effectively reduce the impact on social network data and reduce the loss of information.

According to Fig. 6, with the increase of the number of anonymity, the clustering coefficients of the three algorithms all show a downward trend. When the number of anonymity is different, the clustering coefficients of the anonymous graph constructed by the three algorithms are basically the same as those of the original graph. Among them, the difference between the clustering coefficient of this algorithm

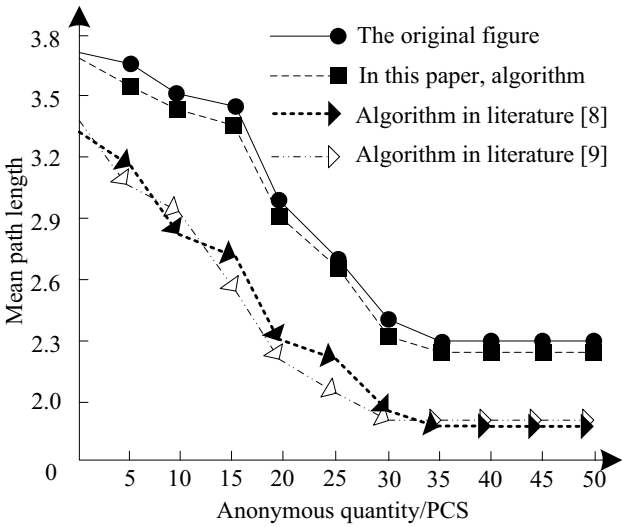

Fig. 5 Test results of average path length

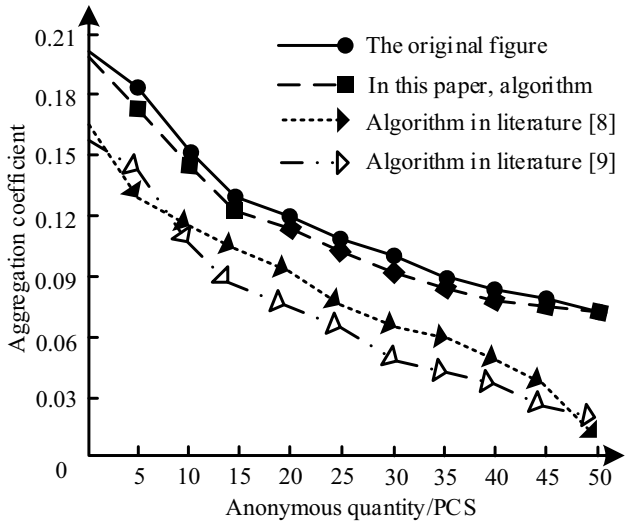

Fig. 6 Test results of aggregation coefficient

and that of the original graph is the smallest, and the clustering coefficients of the other two algorithms are larger than that of the original graph. The experimental results show that the anonymous graph constructed by this algorithm is most similar to the original map in the process of protecting the privacy of end-user social network, which can not only ensure the privacy information security of social network, but also reduce the loss of data utility.

The comprehensive analysis of Figs. 5 and 6, it can be seen that the structural characteristics of the anonymous graph constructed by this algorithm in protecting the privacy of end-user social network is very similar to that of the original map, which shows that this algorithm can not only protect the privacy of end-user social network, but also minimize the loss of social network information.

3.3 End user trajectory privacy protection effect

The three algorithms all cluster the end user track data, so that builds the track anonymous set to realize the end-user track privacy protection. the initial clustering number directly affects the construction effect of the trajectory anonymous set, then further affects the end-user track privacy protection effect. Therefore, the end user track clustering effect of the three algorithms is analyzed when the initial clustering number is different. The end-user track data contains numerical track attributes and classified track attributes, and the track clustering results of these two attributes are shown in Figs. 7, 8 and 9.

According to Figs. 7, 8 and 9, with the increase of the initial clustering number, the better the clustering effect of the three algorithms is. When the initial clustering number is different, the clustering effect of this algorithm is significantly better than that of the other two algorithms. When the initial clustering number is 6, this algorithm can completely separate the end user tracks of the two attributes and gather them together. The other two algorithms can only gather the end user tracks of the two attributes together, and there are still some trajectories that have not yet been clustered. The experimental results show that with the increase of the number of initial clustering, the clustering effects of the three algorithms are improved,

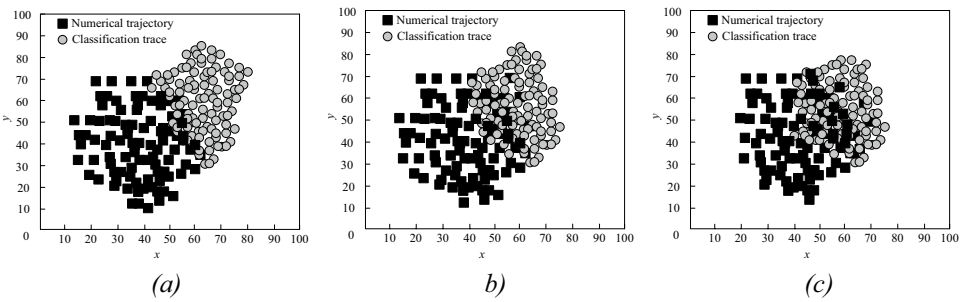

Fig. 7 Clustering effects of the three algorithms when the initial number of clustering is 2 (**a**) Proposed methods (**b**) Methods in Ref.8 (**c**) Methods in Ref.9

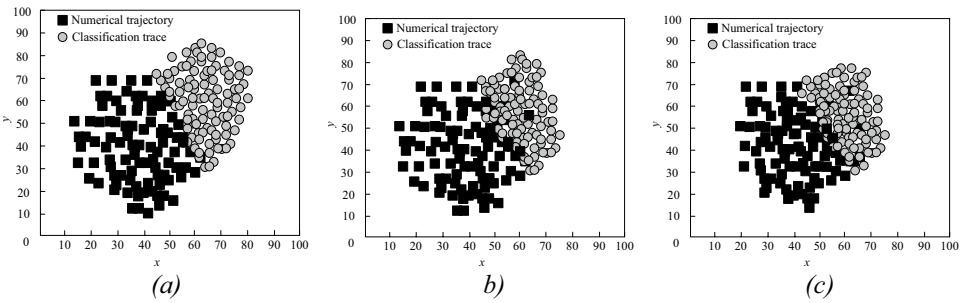

Fig. 8 Clustering effects of the three algorithms when the initial number of clustering is 4 (**a**) Proposed methods (**b**) Methods in Ref.8 (**c**) Methods in Ref.9

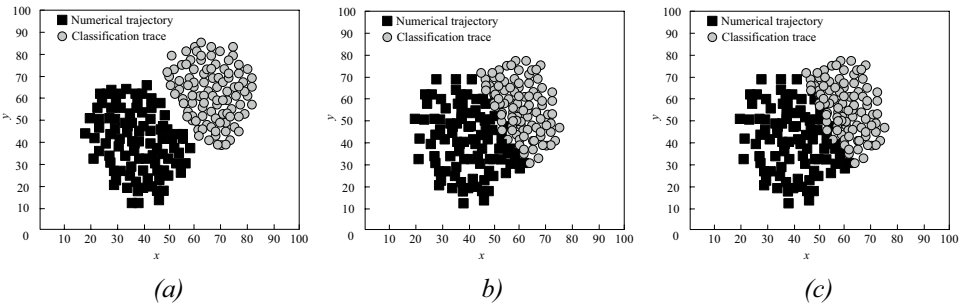

Fig. 9 Clustering effects of the three algorithms when the initial number of clustering is 6 (**a**) Proposed methods (**b**) Methods in Ref.8 (**c**) Methods in Ref.9

and the clustering effect of the algorithm in this paper is the best, which can ensure the best effect of trajectory privacy protection.

In the end user of the intelligent distance education system of the training institution, randomly select an end user, and use the algorithm in this paper to build the trajectory anonymous set for the trajectory of the end user and complete the trajectory privacy protection of the end user. The result of the trajectory privacy protection of the end user is shown in Fig. 10.

According to Fig. 10, the algorithm in this paper can effectively hide the trajectory of the end user and ensure that the privacy of the trajectory of the end user is protected. after dealing with the trajectory is anonymously processed by the algorithm in this paper, there is only a small amount of location information of the end user in the map. And the real trajectory of the user can not be known. Experiments show that the proposed algorithm in this paper can effectively protect the trajectory privacy of end users.

In the case of different trajectory similarity ratios, the relationship between the track loss rate and the privacy protection strength in the process of protecting the trajectory privacy of end user, and the relationship between the proportion of insecure anonymous sets and the privacy protection strength are tested. the test results are shown in Figs. 11 and 12.

According to Fig. 11, keeping the privacy protection intensity consistent, when the end user trajectory similarity ratio is higher, the fewer tracks need to be removed in the generated trajectory anonymous sets of the proposed algorithm, which means the trajectory loss rate is lower than other methods. When the trajectory similarity ratio is consistent, the trajectory loss rate of the proposed algorithm has a positive correlation with the intensity of privacy protection. When the privacy protection strength reaches 8, the trajectory loss rate of different trajectory similarity ratio tends to be stable, ranging from 0.005 to 0.007, all of which do not exceed 0.01.

The experimental results show that the proposed algorithm only needs to remove fewer trajectories to establish the track anonymous set than other methods when protecting the trajectory privacy of end users, which also means that the proposed method effectively reduce the data

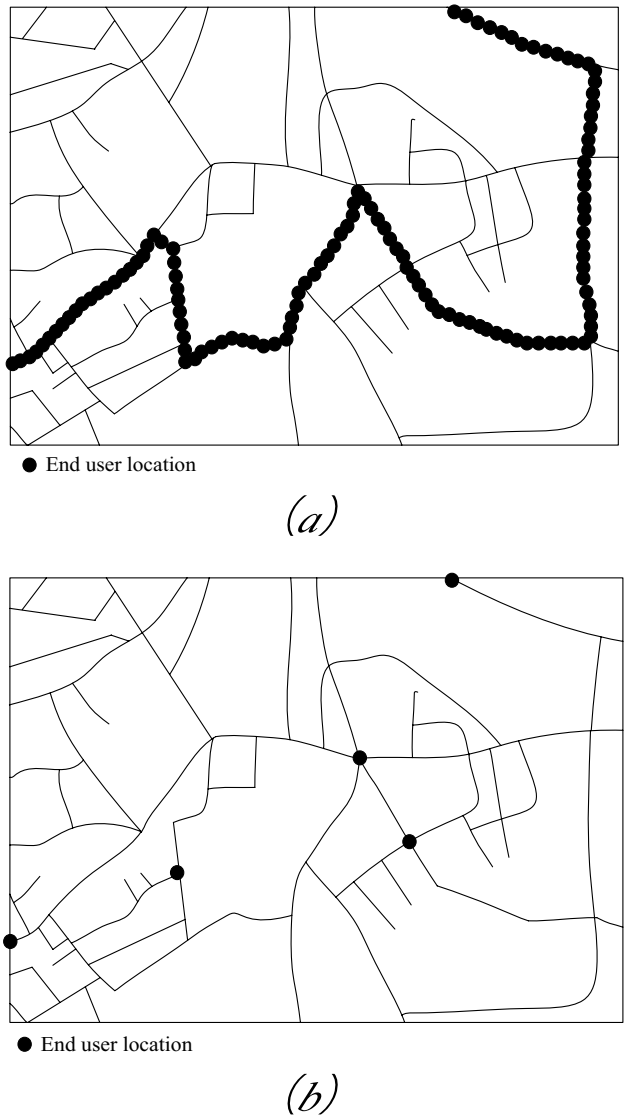

Fig. 10 End user trajectory privacy protection effect of the proposed algorithm (**a**) Original end user trace diagram (**b**)The end user trajectory graph protected by the algorithm in this paper

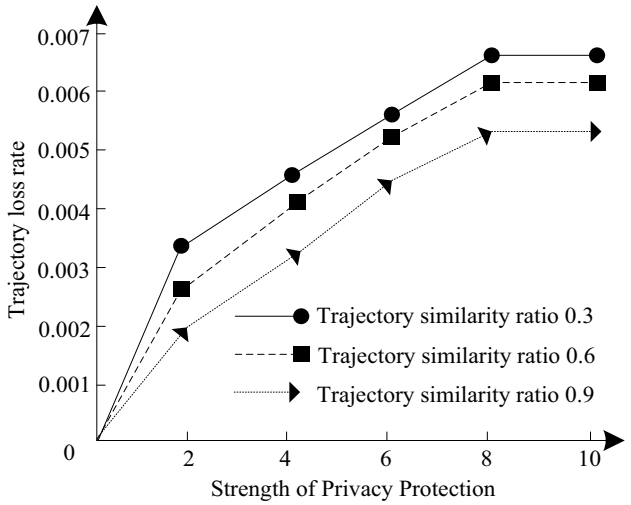

Fig. 11 Relationship between trajectory loss rate and privacy protection intensity

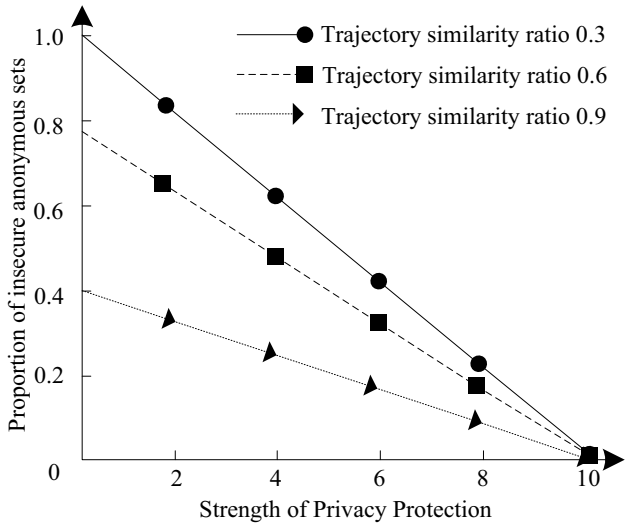

Fig. 12 Relationship between the proportion of insecure anonymous sets and the strength of privacy protection

loss. In fact, the higher the trajectory similarity ratio is, the smaller the data loss will reach when protecting the trajectory privacy of end users.

According to Fig. 12, keeping the privacy protection intensity consistent, when the end user trajectory similarity ratio is higher, the smaller the proportion of insecure anonymous sets of the proposed algorithm has. With the improvement of privacy protection, the proportion of insecure anonymous sets of the three trajectories similar ratios all shows a linear downward trend, and the intensity of privacy protection is inversely proportional to the proportion of insecure anonymous sets. Among them, when the trajectory similarity rate is 0.9, the proportion of insecure anonymous set decreases the slowest; when the protection intensity is 2, its value is 0.35; other similarity values are 0.67 and 0.86, both higher than 0.9. The experimental results show that the higher the trajectory similarity ratio is, the smaller the proportion of insecure anonymous sets of the proposed algorithm is, the better the trajectory privacy protection effect is.

4 Conclusion

The intelligent distance education system can realize the connection between the members of remote working group and users by available wired or wireless networks, so as to achieve the purpose of suspending classes but learning continues during the epidemic. However, the largest problem is the security of end educational users, for this reason, this paper studies the end user privacy protection algorithm of intelligent distance education system, when the privacy protection intensity reaches 8, the trajectory loss rate of different trajectory similarity ratios tends to be stable, ranging from 0.005 to 0.007, all of which are less than 0.01. When the initial cluster number is 6, the algorithm can completely separate and aggregate the end user trajectory of two attributes. Therefore, the methods in this paper include, social network privacy and trajectory privacy to ensure the security of end user privacy.

This paper does not consider the dynamic nature of the social network, which means that the privacy protection technology for the dynamic social network can be introduced to further improve the privacy protection effect of the end user social network in the future.

Acknowledgements The work is sponsored by Natural Science Foundation of Hunan Province with No.2020JJ4434 and 2020JJ5368; Key Scientific Research Projects of Department of Education of Hunan Province with No.19A312; Key Research Project on Degree and Graduate Education Reform of Hunan Province with No.2020JGZD025; National Social Science Foundation of China with No.AEA200013; Industry-Academic Cooperation Foundation of the Ministry of Education of China with No.HKEDU-CK-20200413-129.

Declarations

Ethics declarations The authors have no relevant financial or non-financial interests to disclose. Yating Li and Jiawen Zhu provided the algorithm and experimental results, wrote the manuscript together; Weina Fu revised the paper, supervised and analyzed the experiment. We also declare that data availability and ethics approval is not applicable in this paper.

References

1. Anil KV, Aman S, Edwin L et al (2021) Multilayered-quality education ecosystem (MQEE): an intelligent education modal for sustainable quality education. Journal of Computing High Education 33(3):551–579
2. Santiago IP, Ngel HG, Julián CP et al (2021) Emergency remote teaching and students' academic performance in higher education during the covid-19 pandemic: a case study. Computers in Human Behavior 119(3):106713
3. Choi H, Chung SY, Ko J (2021) Rethinking teacher education policy in ict: lessons from emergency remote teaching (ERT) during the covid-19 pandemic period in Korea. Sustainability 13(10):5480
4. Youcef D, Asma B, Gautam S et al (2022) Fast and Accurate Deep Learning Framework for Secure Fault Diagnosis in the Industrial Internet of Things. IEEE Internet of Things Journal, online first,. https://doi.org/10.1109/JIOT.2021.3092275
5. Shuai L (2019) Introduction of Key Problems in Long-Distance Learning and Training. Mobile Networks and Applications 24(1):1–4
6. Gao P, Li J, Liu S (2021) An Introduction to Key Technology in Artificial Intelligence and big Data Driven e-Learning and e-Education. Mobile Networks & Applications 26(5):2123–2126
7. Seshadri K, Liu P, Koes DR (2020) The 3dmol.js learning environment: a classroom response system for 3d chemical structures. Journal of Chemical Education 97(10):3872–3876
8. Luo B, Li X, Weng J et al (2020) Blockchain enabled trust-based location privacy protection scheme in VANET. IEEE Trans Veh Technol 69(2):2034–2048
9. Wang H, Han G, Zhang W et al (2019) A probabilistic source location privacy protection scheme in wireless sensor networks. IEEE Trans Veh Technol 68(6):5917–5927
10. Wu GF, Tao RX, Yin YS (2019) Social Network Information Transmission Privacy Protection Simulation in the Era of Big Data. Computer Simulation 36(04):107-110+189
11. Chang Y, Zhang SB, Wan GG et al (2019) Practical two-way QKD-based quantum private query with better performance in user privacy. Int J Theor Phys 58(7):2069–2080
12. Liu S, Wang S, Liu X, Gandomi AH, Daneshmand M, Muhammad K, Albuquerque VHC (2021) Human Memory Update Strategy: A Multi-Layer Template Update Mechanism for Remote Visual Monitoring. IEEE Trans Multimedia 23:2188–2198
13. Chun-Wei J, Lin; Gautam Srivastava; Yuyu Zhang, et al (2021) Privacy-Preserving Multiobjective Sanitization Model in 6G IoT Environments. IEEE Internet Things J 8(7):5340–5349
14. Youke W, Haiyang H, Ningyun W et al (2020) An incentive-based protection and recovery strategy for secure big data in social networks. Inf Sci 508:79–91
15. Jimmy MTW, Gautam S, Alireza J et al (2021) Security and Privacy in Shared HitLCPS Using a GA-Based Multiple-Threshold Sanitization Model. IEEE Transactions on Emerging Topics in Computational Intelligence, online first,. https://doi.org/10.1109/TETCI.2020.3032701
16. Liu S, Liu X, Wang S, Khan M (2021) Fuzzy-Aided Solution for Out-of-View Challenge in Visual Tracking under IoT Assisted Complex Environment. Neural Comput Appl 33(4):1055–1065
17. Benke KK, Arslan J (2020) Deep learning algorithms and the protection of data privacy. JAMA Ophthalmology 138(10):1024–1025
18. Asma B, Youcef D, Gautam S et al (2021) A Two-Phase Anomaly Detection Model for Secure Intelligent Transportation Ride-Hailing Trajectories. IEEE Trans Intell Transp Syst 22(7):4496–4506
19. Akremi A, Rouached M (2021) A comprehensive and holistic knowledge model for cloud privacy protection. J Supercomput 6:1–33
20. Shuai L, Dongye L, Gautam S et al (2021) Overview and methods of correlation filter algorithms in object tracking. Complex & Intelligent Systems 7:1895–1917