Jestr

Journal of Engineering Science and Technology Review 15 (2) (2022) 191 - 197

JOURNAL OF
Engineering Science and Technology Review

Research Article

www.jestr.org

Privacy Preservation Method Based on Clustering Interference Algorithm in Social Networks

Ran Zhang[1] and Xianping Wu[2,*]

[1]*College of Mathematics and Computer Science, Tongling University, Tongling 244000, China*
[2]*The Catholic University of Korea, Bu'cheon 100-744, Republic of Korea*

Received 1 March 2022; Accepted 10 May 2022

Abstract

Social network data are very important because they have become an indispensable tool in people's daily life. Therefore, data leakage greatly impacts people. To protect personal privacy information more effectively, a clustering interference algorithm was proposed. The properties of fixed points with the same connectivity were randomly exchanged in the social network while maintaining the entire social network structure by introducing a clustering interference strategy (GIS). The attacker to search for false targets was also induced to protect personal sensitive information data. The information loss in social network structure was analysed and compared by using the MASN (Masking Algorithm for Social Networks) MASN, SaNGreeA (Social Network Greedy Anonymization), and *p*-sensitive *k*-anonymity model. Results show that the privacy preservation method based on the clustering interference algorithm (λ=0.5) has lower normalized information loss at parameters (*k*=2-22) than MASN, SaNGreeA, and *p*-sensitive *k*-anonymity model. In the case of different numbers of quasi-identifiers (0-7), the privacy preservation method based on the clustering interference algorithm (λ=0.5) has the lowest loss of normalized information. Conclusions have practical significance to improve the protection of personal sensitive information in the social network.

Keywords: privacy preservation, social network, clustering, information loss

1. Introduction

With the rapid development of the social network, the online social network has become an indispensable tool in the daily life of people; it has a profound impact on people's social life and social behavior. Considering online social networks with hundreds of millions of users, namely, Wechat, QQ, Sina Weibo, Facebook, Instagram, and Twitter has become the preferred approach to social networking for users. Various activities of users on online social networks generate a large amount of data information, and researchers use information from social networks, such as social network analysis, Sybil defense, and big data processing recommendation system [1], to mine user preferences and other key information from massive data in the social network. However, this type of social network system suffers from data interception, information fraud, privacy spies, and other threats. Releasing social network data is very difficult due to the disclosure of confidential information. It can also mine the identities of many individual entities by analyzing network structure even though some identification data, such as name and social secret data are hidden or deleted from the original data. Therefore, personally sensitive information should be protected before analyzing data in the social network. The main research topic of data analysis is protecting the privacy of the social network.

The privacy preservation method based on social networks shall satisfy the maintainability and privacy of the disclosure of information. Actually, the social network is a collection of multiple social individuals and their related attributes. Social network describes their relationship with individuals and social groups. Social individual attributes are represented as vertices, and the relationship between individuals represents edges that connect with vertices. Privacy attacks may occur in the vertices and edges of the social network. The attributes of vertices in the social network are very important and often contain some identity information, such as ID number, address, telephone, address, and postal code, which can often disclose identifiable information of individuals to varying degrees. Once privacy attackers recognized and identify information, sensitive attributes are exposed. In addition, even if the attacker does not know the attribute, inferring personnel information by obtaining the identities of other active people in the same community is possible. Therefore, privacy preservation should focus on vertices and structure in the social network.

Currently, privacy preservation methods in the social network have been widely studied. Many data privacy preservation and desensitization methods, such as the *k*-anonymity algorithm, *l*-diversity algorithm [2], noise addition, and data disturbance method [3-4], (*a*,*k*)-anonymity, and some improved methods have been widely proposed. Anonymous privacy information protection method and identity attributes are used to further improve data validity under privacy preservation [5]. The protection of information privacy in the social network is still in its initial stage, and practical methods have not yet been developed. At present, most studies focus on how to retain effective community structure or vertex attributes, and few studies consider structure and data information in research on privacy preservation in the social network. Many scholars have been working on establishing model or framework of privacy preservation, which can be controlled by their authorized

*E-mail address: wuxianping@zjcst.edu.cn
ISSN: 1791-2377
doi:10.25103/jestr.152.22

personnel. In addition, the *k*-anonymity algorithm is still very important; it can be integrated into privacy preservation against vertex attackers through neighborhood information in the social network [6]. It uses *k*-anonymity and *l*-diversity for clustering methods of privacy preservation in the social network. Considering the special property of the social network, its privacy preservation methods are different from traditional privacy preservation methods, including some special methods, such as clustering methods, bidirectional graph method [7], dense subgraph mining algorithm [8], and data migration method [9]. The three-level privacy preservation in social network considers attacks by identifying sensitive information on vertices.

Although the privacy preservation algorithm of user information or network structure in the social network has been developed to a certain extent, it is limited to protecting the privacy of network vertices and structural attributes. If the overall structure of the social network changes, then it is unsuitable for community detection. Therefore, this study proposes a social network privacy preservation algorithm based on clustering interference. The proposed algorithm can prevent the privacy and relationship of users from being attacked by local disturbance while maintaining the overall structure. The results show that the algorithm has more advantages in privacy preservation and effectiveness than other algorithms.

This study initially summarizes the research on the privacy model and greedy anonymity method in social network and analyzes the attacks and anonymity in the social network. On this basis, a privacy preservation method based on a clustering interference algorithm was proposed to improve the ability of privacy preservation in the social network.

The remainder of this study is structured as follows: Section 2 reviews the relevant literature from the privacy model in social network, SaNGreeA method, attack in social network, and identification of anonymity in the social network. Section 3 uses the standard "Adult Dataset" composed of vertex and edge structure in social network structure to verify the effectiveness of the algorithm and compares it with other related algorithms. The last section summarizes and concludes this study.

2. State of the Art

2.1 Privacy model in social network

In the social network, attackers can use various information to launch re-identification attacks, and social network users face information security threats and privacy violations [10-11]. Evidently, it cannot fully protect the privacy of users without considering descriptive or structural information. Castiglion et al. [12] proposed a differential privacy preservation scheme based on logical features, and Campan [13] et al. proposed SaNGreeA (SocialNetwork Greedy Anonymization), which protects identifier attributes and structural information from re-identification attacks, without considering sensitive attributes. Assuming that a company publishes attribute information (as shown in Table 1) and social relations (as shown in Fig 1(a)) of employees, we can obtain generalized data with SNGA (as shown in Table 2 and Fig 1 (b)), without explicitly identifying their name. However, if the attacker knows some information, for example, he knows that an employee is 26 years old, then he can easily find out that the salary of the employee is 8000 CNY, which is sensitive to individuals. Yuan [14] proposed k-degree-l-diversity (KDLD) anonymity model to resist degree attack, which is a sensitive property that should satisfy *l*-diversity. However, the method did not consider insensitive properties.

Table 1. Initial tuple of employees' attribute information

Node	Age	No.	Gender	Salary
x^1	25	410076	Male	8000
x^2	25	410075	Male	8000
x^3	27	410077	Male	8000
x^4	35	410099	Male	7000
x^5	38	48201	Female	7000
x^6	36	41075	Female	5000
x^7	30	41099	Male	7000
x^8	28	41099	Male	7000
x^9	33	41075	Female	6000

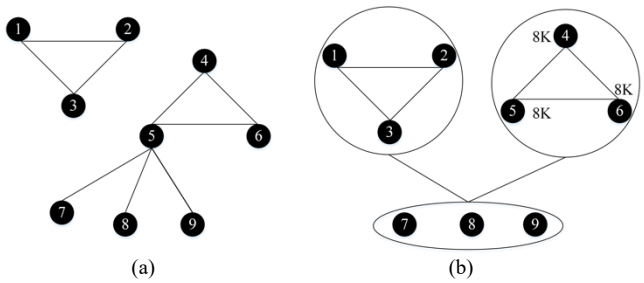

Fig.1. Social Network of Company Staff Information

Table 2. Attribute information generation tuple of company staff

Node	Age	No.	Gender	Salary
x^1	25-27	410***	Male	8000
x^2	25-27	410***	Male	8000
x^3	25-27	410***	Male	8000
x^4	35-38	*****	*	7000
x^5	35-38	*****	*	7000
x^6	35-38	*****	*	5000
x^7	28-33	410***	*	7000
x^8	28-33	410***	*	7000
x^9	28-33	410***	*	6000

The social network model is presented as an undirected graph, with each user represented by vertices and described by tuples containing sensitive attributes and quasi-identifier attributes. The quasi-identifier attributes include numerical and categorical attributes. The categorical attributes are often described as a hierarchical structure; for example, "number" attribute in Table 1 can be expressed as follows: {****}←{{410**}, {482**}}←{{41075}, {41076}, {41099}, {48201}}. The relations between users are represented by edges.

Definition 1 (Social Networks). A social network (SN) is defined as a 3D array G (V, E, A), of which V is a set of vertices, and each vertex represents an individual user. $E \subseteq V \times V$ is a set of edges between two vertices. A is a collection of corresponding tuples of vertices, including sensitive attributes and quasi-identifier attributes.

Definition 2 (Masked Social Networks). Given an initialized social network, SaNGreeA can generate *k*-anonymous masked social networks (MSN), which have quasi-identifier and relation attributes. MSN is defined as a

3D array, where $MV = \{cl_1, cl_2, ..., cl_m\}$ is the division of edges in social networks, including "*m*" clusters, can be generated by SNGA (Social Network Greedy Anonymization). Each cluster contains at least "*k*" vertices, and $cl_i \cap cl_j = \varnothing (i \neq j)$, $\bigcup_{i=1}^{m} cl_i = V$. *ME* represents the subset of edges between clusters: $(cl_i, cl_j) \in ME(i \neq j)$, if $cl_i \cap cl_j = \varnothing (i \neq j)$, and $\bigcup_{i=1}^{m} cl_i = V$.

2.2 SaNGreeA Method

SaNGreeA is a greedy clustering algorithm that masks the initial social networks to generate a *k*-anonymous network, including the generalization of quasi-identifier attributes and relations. According to the measurement of information loss, all vertices of the initial social network are divided into multiple clusters. Each cluster in the masked social network should contain *k* vertices at least. The generalization of quasi-identifier attributes and relations reduces data quality. Each time, SNGA selects the vertex that generates a minimal loss of information and adds it to the current cluster. Once the cluster satisfies *k*-anonymity, a new cluster is created until the entire network is masked.

Definition 3 (Loss of Generalized Information), pp. *cl* is a cluster, and its quasi-identifier attribute is $QID = \{N_1, ...N_p, C_1, ..., C_q\}$, of which $N_i (i \in [1, p])$ is a numeric attribute, and $C_j (j \in [1, q])$ is a classification attribute. The loss of generalized information about clusters *cl* is formed by *QID*, as follows:

$$GIL(cl) = |cl| \cdot \left(\sum_{i=1}^{p} \frac{\max_{cl}(T[N_i]) - -\min_{T \in A_{cl}}(T[N_i])}{\max_{T \in A}(T[N_i]) - \min_{T \in A}(T[N_i])} + \sum_{i=1}^{q} \frac{h(H_{cl}[C_i])}{h(H[C_i])} \right) \quad (1)$$

where *T* is a tuple, *A* is the tuple set in the initial graph, *H* is the hierarchical set of categorical attribute values in the initial graph, and h is the height of the hierarchical structure.

Definition 4 (Normalization of generalized information loss). Based on division $S = \{cl_1, ..., cl_n\}$ and masked graph *G*, the normalization of generalized information loss is obtained as follows:

$$NGIL(G, S) = \frac{\sum_{i=1}^{n} GIL(cl_i)}{|V| \cdot (p + q)} \quad (2)$$

Definition 5 (Distance). The distance between vertex *V* and cluster *cl* is obtained as follows:

$$D(v, cl) = \frac{\sum_{v} \left| \{ v_x | v_x \in V, v_x \neq v, v_x \neq v^*, R(v_x, v) \neq R(v_x, v^k) \} \right|}{\left| c \right| \cdot (|V| - 2)} \quad (3)$$

where $R(v_1, v_2)$ indicates that v1 and v_2 are friend nodes.

2.3 Attack analysis in social networks

Various knowledge and means can be used to attack privacy given the diversity of social networks [15-16], thereby causing a great challenge to the protection of private information in social networks. Privacy disclosure in social networks may occur in vertex attribute, the relationship between vertices, network structure, node structure in social network, and importance of user location. The attack on network structure is one of the most popular privacy attacks in social network privacy preservation based on known topological information by comparing local structure before and after data release, which can identify the identity of users. The background knowledge of the attacker is divided into a graph structure, node information, edge information, and prediction model. A series of attacks can obtain the specific anonymity of vertices in social network to check whether a link exists between vertex pairs, including active attacks and passive attacks on anonymous social networks [17]. Active attacks need to create some new vertices or links. An attacker selects any set of recorders that need to violate privacy in an active attack. It creates some new vertices linked to targeting vertices and builds the link mode between these new vertices to enable them to stand out from the anonymous network structure. Passive attacks simply attempt to discover their release in the network and relations between target vertices. Vertex alliances can destroy the privacy of adjacent vertices by passive attacks.

The structure of the network itself decides the degree to which individuals can be distinguished in social networks, and attackers can obtain relevant information based on detailed queries for vertex and subnet information surveys. The structural attack model of network information is a type of attack. Attackers can select a set of vertices in some way and express the trust relationship between them. For attacks on *k*-neighbor maps, attackers have maps from specific vertices and d-hop from its community. Another attack assumes that attackers can select a large number of subnets to partially overlap with the original network.

2.4 Identification of anonymity in social networks

The privacy preservation strategy in the social network is divided into anonymous vertex, anonymous subgraph, data interruption, and clustering. Anonymous vertex ensures that the probability of identifying vertex information is less than $1/k$ at least through improved social networks. Data interference aims to prevent the inference of raw data by randomly modifying the network, which can be divided into noise addition and noise interference. The idea of noise addition takes the form of a random addition to add noise to social networks, thereby removing or switching to opponent links to identify target vertices in the network and their relations. Stergiou et al. [18] designed a spectrum-protected randomization method that can better preserve functions of the network by adding or removing a spectrum-switching design without sacrificing considerable privacy preservation in the process of randomization. Noise interference is achieved by adding, modifying, or deleting links to create new *k*-anonymous network, in which each vertex has at least "*k*" other vertices of the same degree. Anonymous network uses generalized vertex labels and inserted links until no neighborhood can be distinguished at probability $1/k$. Although these methods can play a role in preventing the re-identification of attacks on network structure, some aspects reflect error import, node interruption, and topology [19-20].

3. Methodology

3.1 Social network structure information

Sensitive information that concern attackers can usually be located in local communities for the large-scale social

network, and privacy preservation operations are required to minimize changes in the overall structure of the network to maintain data availability. In addition, social networks are usually sparse, and most direct connections between individuals are limited. Therefore, community division is integrated into the proposed algorithm.

Structural similarity can be used to measure the local connectivity density of any adjacent vertices in an undirected network, considering the network was described as a set $G=(V, E)$, the structural similarity sij between any two adjacent vertices i and j is as follows:

$$S_{ij}=\frac{e_{ij}+\Gamma(i)\cap\Gamma(i)}{k_i+k_j} \tag{4}$$

where $\Gamma(i)$ and k_i represent the degree of neighborhood set and verte x_i, respectively. If vertex i is connected with vertex j, then $e_{ij}=1$; otherwise $e_{ij}=0$. By applying structural similarity, the local structure measurement standard can be described as follows:

$$T(c)=\frac{S_c^{in}}{S_c^{in}+S_c^{out}} \tag{5}$$

where $S_c^{in}=\sum_{i,j\in c\wedge e_{ij}\in E} S_{ij}$ is the internal similarity c of the community, and $S_c^{out}=\sum_{i\in c\wedge j\notin c\wedge e_{ij}\in E} S_{ij}$ is the external similarity. *T(c)* can be used to measure the quality c of specific community. Formula (5) is used to determine when the variable *T* in the community is great, the vertex connection becomes denser, and the connection between different communities becomes sparser. This trend is consistent with the definition of the community. When the adjacent vertex *i* is added in community *c*, the local structure measurement standard of *c* is changed, and the increment of its community c can be expressed in the following format (6) after vertex *i* is merged into *c*:

$$\Delta T_c(i)=T(c\cup\{i\})-T(c) \tag{6}$$

3.2 Description of algorithm

This study proposes a clustering interference algorithm based on social network privacy preservation. The proposed method improves the ability of privacy preservation in social network according to the power-law distribution of vertex features and local clustering structure information in the social network. The algorithm initially performs the information exchange of nodes with the same connectivity, and then the social network is divided into some clustering sets by using minimal data loss and community structure loss. Finally, the interference strategy includes adding, deleting, and modifying internal and external node edges, which can enhance the structural anonymity of each cluster and large vertex to avoid destroying the extension of the local community structure. The edge adding and deleting must be symmetrical and complementary.

Algorithm 1: Clustering Interference Algorithm

Input：Graph intmv=(max(degree)+min(degree))/2 , round figure *k*;
Output: Graph $G=(V', E')$;
Remove all identity information from original graph to generate the Graph *G*;
Build adjacent matrix g*M* atri *x*[*n*][*n*] and digit group degree[*n*];
intmv = (max(degree) + min(degree)) / 2
Construct a vector *V*[*m*], where *m* refers to the number of different values of digit group degree[*n*], and each element is a vertex vector with the same vertex connectivity.
for d=0:m
Exchange all attribute information for any two unconnected vertices in vector *V*[*m*].
end for
$NACA(k,V,gMatrix,\lambda)$
$c\notin C$

Exchanging information about these vertices in the whole network helps to induce false attack targets and prevent structural re-identification. At the same time, the exchange information between vertices of the same degrees does not change the topology in the original social network. In essence, vertex exchange is only an attribute exchange rather than topological attribute. Network Anonymous Clustering Algorithm (NACA) consists of partition clustering and adjustment clustering, as shown in Algorithm 2.

Algorithm 2: $NACA(k,V,gMatrix,\lambda)$

$C=\phi; i=1; H=\phi; C^{tmp}=\phi; G=\phi; E^c=\phi$;
While ($|V|\neq 0$) {
Choose a vertex as Xseed with maximum connectivity from V-C
$C_{i++}=X_{seed}$;
$V=V-X_{seed}$;
H=SearchNeighbors(C_i C_i);
While($|Ci|<k\wedge|H|\neq 0$);
$C_{tmp}=C_i\cup\{X^*\}; X^*\in H; G^*=\{X|X\in C_{tmp}\}$;
Conclude vertex attribute from Ctmp;
$X^*=\arg\min_{X^*\in H}((1-\lambda)\cdot NTQL(G^*,C_{tmp})+\lambda\cdot\tau_c(X^*))$;
$C_i=C_i\cup\{X^*\}; V=V-\{X^*\}$;
SearchNeighbors();
};
$C=ACluster(C_i,C,k)$; {
if ($|C|>1$);
for ($x=1; x<|C|; x++$), $E^c=E^c\cup\{E_{ix}^c\}$;
};

NACA, the initial node of the community, is the maximum degree of the unselected node-set. According to the adjacency of the network, the neighbor set is searched from the selected node by the function SearchNeighbors(). The node can join the current community within minimal local structure loss and information loss. According to the change in minimal data loss and community structure, to strengthen the anonymity of social network structure, a novel CIA is proposed, as shown in Algorithm 3. The complementary combination of adding, deleting, and modifying the clustering interference edges is performed on large vertices of each community and the boundary between two communities.

Algorithm 3: $GIS(C,E^{inter},V^{inter},V,mv)$

$V^{big}=\phi; A^c=\phi; E^c=\phi$;
for each $e_{ij}\in E^{inter}\wedge v_i,v_j\in V^{inter}\wedge v_i\in C_m\wedge v_j\in C_n\wedge m\neq n$ do ;
$vx=\arg\min_{x\in C_m\wedge x\notin V^{inter}}|\deg ree[v_i]-\deg ree[v_x]|; v_y=\arg\min$;

$e_{xy} = q; e_{ij} = 0$;
end for;
for each cluster $c \in C$ do;
$V^{big} = SearchBig\ Vertices(c, mv)$;

2.3 Experiment setting

The experimental dataset uses the standard adult dataset composed of vertices and edge structures in the social network, and this dataset is used to study the working mechanism and sensitivity of the proposed network privacy preservation algorithm to these parameter settings involving λ and k. SaNGreeA [21], MASN (Masking Algorithm for Social Networks), and p-sensitive k-anonymity model [22] are used to analyze attribute information loss and network structure information loss under different parameters and different numbers of quasi-identifiers. The software and hardware environments are shown in Table 3.

Table 3. List of experimental environment details

Item	Details
Processor	Intel(R) Core(TM) i7-6700H
RAM	24GB
Operating system (OS)	Windows10 (64-bit)
Video Card	NVIDIA GeForce GTX 980M (8192MB)

4. Results Analysis

Information loss and characteristic statistics in social networks are usually used to evaluate the performance and effectiveness of privacy preservation algorithms in social networks. Generally, more information loss indicates lower data availability and lower effectiveness of the algorithm. On the contrary, less information loss indicates higher data availability and a more acceptable algorithm. Moreover, the smaller difference in statistical characteristics between the original network and the published network indicates more efficient algorithm.

The statistics of information loss of several models are shown in Fig 2. This study implements MASN, SaNGreea, p-sensitive k-anonymity mode, and clustering interference algorithm based on the setting of different k parameters. The range of parameter K is 0-22, and the number of quasi-identifiers is 4. The parameter λ in the clustering interference algorithm is used to adjust the ratio of normal data information loss and structure information loss.

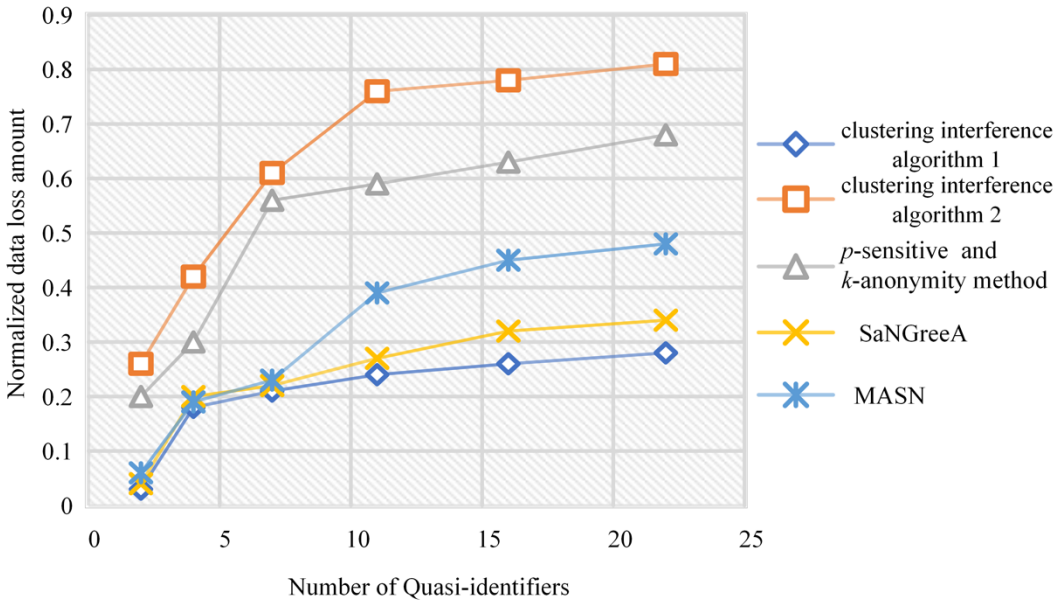

Fig. 2. Information loss of several algorithms under different parameters

The experimental results show that the normalized information loss of the clustering interference algorithm (λ =0.5) in the parameter (k=2-22) is lower than that of MASN, SaNGreeA, p-sensitive k-anonymity mode.

This study also sets the information loss statistics of several models under different numbers of quasi-identifiers. The experimental results are shown in Fig 3, and the value range of the number of quasi-identifiers is 0-7, which indicates that the normalized information loss of the clustering interference algorithm (λ =0.5) is the lowest. For almost all privacy preservation algorithms, a common feature is that information loss increases with the increase in the number of quasi-identifiers, suggesting that selecting an appropriate number of quasi-identifiers is beneficial for reducing information loss of privacy preservation. The number of quasi-identifiers is an inherent feature of the dataset, but some attributes with lower importance can be removed to reduce the number of quasi-identifiers and improve the effectiveness of privacy preservation. In the local region, the correlation of attributes is evident. The local community divided in the proposed algorithm is used to collect correction attributes.

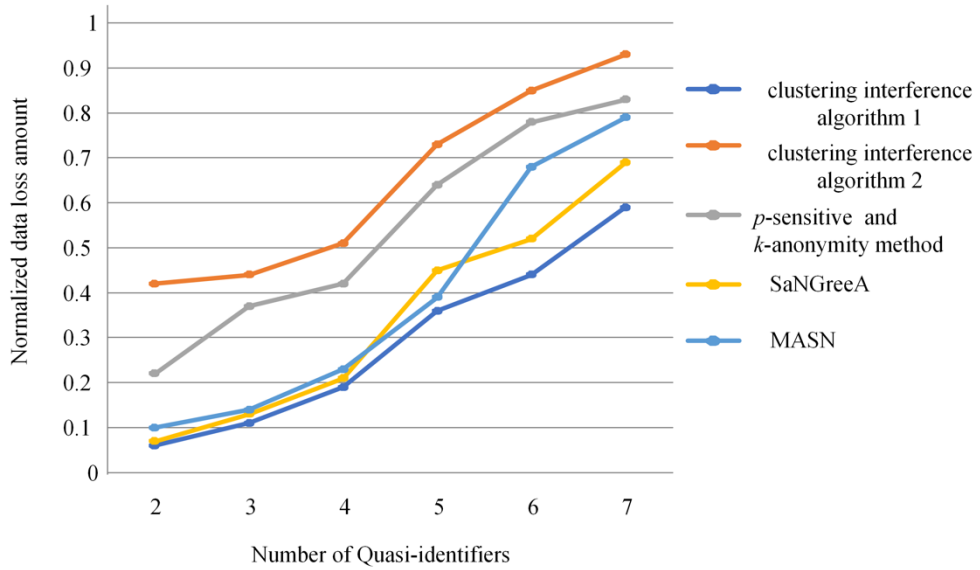

Fig. 3. Information loss of several algorithms under different numbers of quasi-identifiers

5. Conclusion

With respect to the problems encountered in cluster mining of privacy preservation in the social network, a clustering interference algorithm was proposed to protect the privacy of data released in the social network through the analysis of node attributes and structure attributes and satisfy privacy preservation in social networks. The following conclusions are obtained:

(1) The normalized information loss of the clustering interference algorithm (λ=0.5) in the parameter (k=2-22) is lower than that of MASN, SaNGreeA, and p-sensitive k-anonymity mode.

(2) The normalized information loss of the clustering interference algorithm (λ=0.5) is the lowest under different numbers of quasi-identifiers (0-7).

In this study, the proposed algorithm is used in the standard adult dataset without other data, and the generalization ability of the model is tested. Different types of data can be used for exercise in subsequent studies to improve the generalization ability of the model.

Acknowledgment

This study was supported by the Natural Science Research Project of Anhui Province Universities (No. KJ2010B234).

References

1. Vaghashia, H., Ganatra, A. A., "Survey: Privacy preservation techniques in data mining". *International Journal of Computer Applications*, 119(4), 2015, pp.20-26.
2. Zhou, B., Pei, J., "The *k*-anonymity and *l*-diversity approaches for privacy preservation in social networks against neighborhood attacks". *Knowledge and Information Systems*, 28(1), 2011, pp.47-77.
3. Xia, Z., Wang, X., Zhang, L., Qin, Z., Sun, X., Ren, K., "A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing". *IEEE Transactions on Information Forensics and Security*, 11(11), 2016, pp. 2594-2608.
4. Fu, Z., Wu, X., Guan, C., Sun, X., Ren, K., "Toward efficient multi-keyword fuzzy search over encrypted outsourced data with accuracy improvement". *IEEE Transactions on Information Forensics and Security*, 11(12), 2016, pp. 2706-2716.
5. Zhang, Z., Sun, R., Zhao, C., Wang, J., Chang, C. K., Gupta, B. B., "CyVOD: a novel trinity multimedia social network scheme". *Multimedia Tools and Applications*, 76(18), 2017, pp. 18513-18529.
6. Yu, C., Li, J., Li, X., Ren, X., Gupta, B. B., "Four-image encryption scheme based on quaternion Fresnel transform, chaos and computer generated hologram". *Multimedia Tools and Applications*, 77(4), 2018, pp. 4585-4608.
7. Li, J., Li, X., Yang, B., Sun, X., "Segmentation-based image copy-move forgery detection scheme". *IEEE Transactions on Information Forensics and Security*, 10(3), 2014, pp. 507-518.
8. Sapountzi, A., Psannis, K. E., "Social networking data analysis tools & challenges". *Future Generation Computer Systems,* 86, 2018, pp. 893-913.
9. Stergiou, C., Psannis, K. E., "Efficient and secure big data delivery in cloud computing". *Multimedia Tools and Applications*, 76(21), 2017, pp. 22803-22822.
10. Mortazavi, R., Erfani, S. H., "GRAM: an efficient (*k, l*) graph anonymization method". *Expert Systems with Applications*, 153, 2020, pp. 113454.
11. Kim, H., Hovav, A., Han, J., "Protecting intellectual property from in- sider threats: A management information security intelligence perspective". *Journal of Intellectual Capital*, 21(2), 2019, pp.181-202.
12. Castiglioni, V., Chatzikokolakis, K., Palamidessi, C. A., "logical characterization of differential privacy". *Science of Computer Programming*, 188, 2020, pp. 012388.
13. Campan, A., Truta, T. M., "Data and structural k-anonymity in social networks". *International Workshop on Privacy, Security, and Trust in KDD. Springer, Berlin, Heidelberg,* 2008, pp.33-54
14. Yuan, M., Chen, L., Philip, S. Y., Yu, T., "Protecting sensitive labels in social network data anonymization". *IEEE Transactions on Knowledge and Data Engineering*, 25(3), 2011, pp.633-647.
15. Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., Philip, S. Y., "A cpmprehensive survey on graph neural network". *IEEE Transactions on Neural Networks and Learning Systems*, 32(1), 2020, pp. 4-24.
16. Zhang, Z., Cui, P., Zhu, W. "Deep learning on graphs: A survey". *IEEE Transactions on Knowledge and Data Engineering*, 34(1), 2020, pp. 249-270.
17. Gupta, B. B., Arachchilage, N. A., Psannis, K. E., "Defending against phishing attacks: Taxonomy of methods, current issues and future directions". *Telecommunication Systems*, 67(2), 2017, pp.247-267.

18. Stergiou, C., Psannis, K. E., Kim, B. G., Gupta, B., "Secure integration of IoT and cloud computing". *Future Generation Computer Systems*, 78, 2018, pp.964-975.
19. Ji, S., Wang, T., Chen, J., Li, W., Mittal, P., Beyah, R., "De-sag: On the de-anonymization of structure-attribute graph data". *IEEE Transactions on Dependable and Secure Computing*, 16(4), 2017, pp. 594-607.
20. Qian, J., Li, X. Y., Zhang, C., Chen, L., Jung, T., Han, J., "Social network de-anonymization and privacy inference with knowledge graph model". *IEEE Transactions on Dependable and Secure Computing*, 16(4), 2017, pp.679-692.
21. Fu, Z., Ren, K., Shu, J., Sun, X., Huang, F. Enabling personalized search over encrypted outsourced data with efficiency improvement. *IEEE Transactions on Parallel and Distributed Systems*, 27(9), 2015, pp. 2546-2559.
22. Zhou, Z., Wang, Y., Wu, Q. J., Yang, C. N., Sun, X. Effective and efficient global context verification for image copy detection. *IEEE Transactions on Information Forensics and Security*, 12(1), 2016, pp.48-63.