A Privacy-preserving Interactive Messaging Scheme Based on Users Credibility over Online Social Networks

Shunan Zhang [1], Ying Cai *[2], Hongke Xia [3]

Department of Computer Science, Beijing Information Science & Technology University, Beijing, 100101, China

[1]291469736@qq.com

[2]ycai@bistu.edu.cn

[3]kk325@126.com

***Abstract*— Due to the tremendous growth in user population over online social networks (OSNs), interactive messaging traffic has experienced exponential increase, which consequently elevates the privacy leakage. Malicious users may utilize camouflage, phishing and other unconventional techniques to lure ordinary users to reveal their sensitive attributes in their interactive messages, when they least expect that the other sides are malicious. Therefore, it is very important to identify malicious users quickly and successfully when socializing over OSNs. In this paper, we propose a privacy-preserving interactive messaging scheme by relying on user credibility, which characterizes users' reputation through users' social behaviors, and use the item response theory in psychometrics to assess the risk levels of users' interactive messages. We have conducted extensive studies by sending interactive messages with sensitive attributes and have observed that the success rate of identifying malicious users based on user credibility is significantly higher and our scheme can effectively lower the risk of privacy leakage in our interactive messaging system.**

***Index Terms*— Online Social Networks, Interactive Messaging, User Credibility, Item Response Theory, Risk**

I. Introduction

Online Social Networks (OSNs) have become an indispensable part of our daily lives. Users can get to know a large number of new friends through OSNs and obtain more information by exchanging messages. Unfortunately, such convenience comes without price. In the process of socializing, users often disclose their private information intentionally or unintentionally.

Due to national privacy policy restrictions, neither the social platform service provider nor the third party service facilitator has the right to disclose users' interactive contents containing private information. These policies make it difficult to measure privacy in interactive messages, so there is relatively little privacy protection schemes for interactive messages. Some schemes set users' authority of interactive messaging only based on users' attitude towards privacy concerns. Others attempt to protect users' privacy by optimizing user privacy settings. However, the process of manually setting privacy is very cumbersome, which will wear out users' patience and they may just skip the stage of privacy settings. Some schemes suggest using early warning measures, warning users before they send sensitive messages to others in an untrusted network environment. Unfortunately, the actual effect of this early warning mechanism is not really effective because many users tend to ignore the warning and send sensitive messages out without scrutiny.

In this paper, by taking into account the characteristics of users' interactive messages and the disadvantages of existing schemes, we propose a privacy protection scheme based on user credibility. Our scheme measures the user credibility based on users' social behaviors over OSNs. Social behaviors can directly capture the inherent characteristics of a user. For example, we can learn a user's personality from the number of friends he/she often contacts with and the number of interactive messages exchanged. It can also capture the relationships indirectly among users. For instance, we can determine whether two users are friends through the contact frequency. Therefore, a user's credibility can be fine-tuned to determine whether a user is malicious. On the other hand, since the users' interactive messages may contain his/her sensitive attributes, we also need to characterize the risk level for interactive messages. The captured risk level of a user's interactive messages can be used to control whether the user sends messages or not.

The rest of the paper is organized as follows. Section 2 presents the related work. In Section 3, we introduce the adversarial model, a user's credibility and messages risk assessment mechanism, and describe how to measure the credibility and the risk level. In Section 4, we conduct performance evaluation of our scheme and show its effectiveness in addressing privacy concerns in interactive messaging. Section 5 concludes the paper.

II. Related work

The definition of online social networks (OSNs) was first formally proposed by Boyd et al.[1] in 2007. Since then, OSNs have received intensive attention. Wang et al.[2] modeled

*Ying Cai is the corresponding author.

978-1-5386-4502-4/17/$31.00 ©2017 IEEE

social threats from social networking sites, third-party service providers, social network users, and attackers, and considered six security factors. This model can be used to effectively analyze the behaviors of attackers and capture privacy leaks in detail. In face of various privacy leaks, many research works have been carried out.

A. Privacy Score Mechanism

Privacy is a very difficult to address as different users have different standard for their privacy. Privacy can be protected by using privacy setting according to privacy scores. Liu et al.[3] used the item response theory (IRT) to evaluate users' privacy scores, which can be used to recommend privacy settings for users. However, the scheme relies on users' attitude towards sensitive attributes, and hence it depends on a certain degree of subjectivity. Li et al.[4] proposed a trust-based privacy setup system to help users choose privacy preferences. Unfortunately, this setup process of privacy setting still contains users' subjective preference. Srivastava et al.[5] used privacy entropy to measure the privacy degree of user identity messages. Vidyalakshmi et al.[6] designed a privacy score function based on the cubic Bezier curve according to users' privacy preference to determine whether a user can join the circle of friends. Qian et al.[7] constructed a model of privacy reasoning based on attackers' knowledge. However, this model has not considered the correlation of prior knowledge and the characteristics of the corresponding probability distributions. Petkos et al.[8] summarized the privacy scores for privacy settings, including user-defined attribute sensitivity and developed mechanisms to set different levels according to the combination of attribute messages. This indeed provides a theoretical foundation for privacy score research. Otsuki et al.[9] proposed a method for measuring the priority of user attributes and investigated the sensitivity of message attributes, which could capture users' privacy attitude objectively.

B. Trust or Reputation Based Privacy Protection

Privacy can also be protected based on user controlled message sharing according to either trust or reputation. Li et al.[10] used users' trust degree and contact time intervals to measure user credibility, but the definition of trust is too coarse. Abbasi et al.[11] measured the credibility of messages sent by users based on the propagation of error messages and rumors. The measurement took into account the similarity of users' behavior, but failed to specify how to quantify users' behavior. Guo et al.[12] designed a user reputation system according to users' behaviors, which verified users' property by the third party platform and voted for authenticated users based on reputation. Gambhir et al.[13] defined trust factors based on users' behaviors and interactive messages to determine user credibility. Unfortunately, there is lacking sound theoretical foundation for the definition of users' behaviors and the weight of interactive messages.

C. Other Privacy Protection Approaches

There are also other related works addressing privacy issues. Reynaert et al.[14] proposed a framework to improve privacy protection, which was primarily based on the anonymity of social graphs and the security of messages flow in browsers. Palomar et al.[15] developed an access control protocol for message sharing, which could set permissions for data sharing by the data owner and co-managers together. Li et al.[16] designed a node perturbation algorithm based on k-anonymity, which could effectively prevent one hop or multi-hop friends from revealing anonymous messages. Ahmed et al.[17] proposed a method of adding virtual nodes to ensure k-anonymity, which could deal with recognition attack of a group of messages.

III. System Design

In this section, we present our scheme on the privacy protection for interactive messaging over OSNs and elaborate it from three aspects: attack model, user credibility and messages risk assessment mechanism.

A. Attack Model

Depending on interactive messages, attacks can be divided into two categories. First, we assuming that there are an attacker A and an ordinary user U in an OSN site. The friends of U, the family of U, and the colleagues of U are also registered with this OSN site. The attribute set of U is $ATTR_U(attr_1, attr_2, attr_3, attr_4, ...)$. Attacker A obtains some attributes of U previously by some means of attacks and registers with the OSN site as U_A, which has a set of attributes $ATTR_{U_A}(attr_1, attr_3, attr_4, ...)$ that are highly similar to U. Due to different people identify U by different combination of attributes, so they may mistake A for U, and may leak their own privacy to A, as shown in Fig. 1. Then attacker A can obtain their relatives and friends' privacy by the same way after getting their privacy.

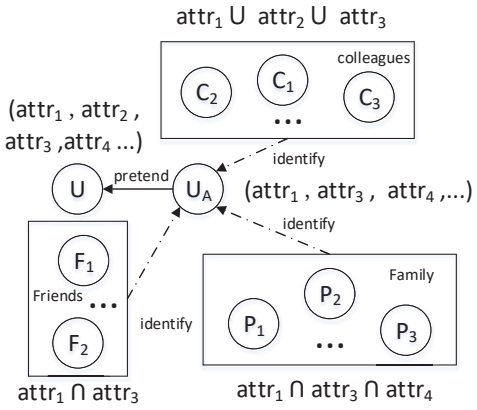

Fig. 1. The first type of attack Model

Second, suppose that there are an attacker A and an ordinary user set $SET_U(U_1, U_2, U_3, ...)$. Attacker A uses the obtained fake attributes to register as U_A. In order to obtain the private information from a larger number of OSNs users, U_A sends friend requests to all users in SET_U. For some users who like to make friends, the request will be granted under normal

circumstances. Once U_A's requests are granted, U_A will push messages with malicious links to users who become friends of U_A in SET_U. As soon as a user clicks to a link, he/she may leak his/her privacy to U_A. When U_A reaches this goal, he could delete the existing friends and then continue the next round of attacks.

B. User Credibility

In order to protect users' privacy in interactive messaging systems, it is important to recommend users to manage their interactions with cautions in terms of whom they interact with and what they exchange. For user to user interactions, users need to add each other as their friends. Unfortunately, many users do not usually verify each other before granting their requests, which introduces a potential threat for privacy invasion.

The aforementioned situation displays the fact that many users have not paid enough attention to protect their privacy, which will potentially result in privacy leakage of many other users. To deal with this problem, we introduce the concept of user credibility, Cre, based on users' social behaviors, as defined in Definition 1 below. By using this metric, we could improve the probability of recognizing malicious users and limit the social interactions with low-confidence users.

Definition 1 (User Credibility): User Credibility, Cre, for a user is defined as the degree that the user is credible to all other users over the OSN considered. It is characterized by the user's social behaviors.

Before presenting our method to mathematically characterize the user credibility in detail, we need to define the precise definition of a completed session between a user and another user as follows.

Definition 2 (Completed Session): A completed session between user A and user B is defined as the interactive session between A and B with at least a prescribed time period of silence, say, T, before and after the session, i.e., the time interval between the last message of this session and the first message of the next session needs to be greater than T, where T depends on the specific OSN environment. In other words, a completed session between A and B is the session from the start of the interaction to the last message followed by a silence period of a prescribed time T.

Now we are ready to give our method to calculate user credibility. We assume that in the interaction process between user A and user B, the factors to be considered in the measurement of $Cre_{U\leftrightarrow U}(A,B)$ can be expressed as $< P_{rep}(A,B), P_{con}(A,B), P_{ses}(A,B) >$. $P_{rep}(A,B)$. $P_{rep}(A,B)$ denotes the message response rate of user B to user A, as shown in equation (1):

$$P_{rep}(A,B) = \frac{T_{rep}(B,A)}{T_{sen}(A,B)} \tag{1}$$

where $T_{rep}(B,A)$ denotes the number of replies of B to A in all sessions which A sends the first message, $T_{sen}(A,B)$ denotes the times A first sends messages to B in all sessions. $P_{con}(A,B)$ denotes the ratio of effective days of interactions between A and B, as shown in equation (2):

$$P_{con}(A,B) = \frac{D_{con}(A,B)}{D_{all}(A,B)} \tag{2}$$

where $D_{con}(A,B)$ denotes the effective days of interactions between A and B, $D_{all}(A,B)$ denotes the total number of days when A and B are friends so far. One effective day of interactions is defined as both sides have at least one reply to each other in one day. In other words, within a day, if A sends a message to B, B should reply to A, and then A should continue reply to B. $P_{ses}(A,B)$ is the effective session ratio between A and B as defined in equation(3):

$$P_{ses}(A,B) = \frac{T_{ses}(A,B)}{T_{all}(A,B)} \tag{3}$$

where $T_{ses}(A,B)$ denotes the number of completed sessions between A and B. $T_{all}(A,B)$ denotes the number of sessions between A and B (including the sessions that one party sends a messages, but does not receive reply).

Due to the frequency of interactions are different between A and each of his/her friends, we assume that the friends with $P_{rep}(A,B)$, $P_{con}(A,B)$ and $P_{ses}(A,B)$ equal to 0 are negative friends, while the rest of the friends were active friends. We only calculate A's credibility through positive friends. Meanwhile, user A needs to maintain two lists of friends. One is the list of current friends, which contains the existing friends of user A so far (not including deleted friends) and the number is denoted as l. The other one is the list of past friends, which contains all friends who have been added by user A since A creates the account (including deleted friends) and the number is denoted as m ($l \in m$). The number of active friends of user A is n, then the A's credibility $Cre(A)$ can be calculated as in equation (4):

$$\begin{aligned} Cre(A) = (\frac{1}{n}\sum_{i=0}^{n} a{\cdot}P_{rep}(A,B) + b{\cdot}P_{con}(A,B) \\ +c{\cdot}P_{ses}(A,B)){\cdot}(\frac{n}{l}+\frac{l}{m}){\cdot}\frac{1}{2} \end{aligned} \tag{4}$$

where n/l is the proportion of active friends. When n is larger, it may indicate that the popularity of the current user in OSNs is higher. l/m reflects the purpose of the current user's social activities. The smaller the ratio is, indicating that the user has deleted more friends, the higher likelihood that it is a malicious user. a, b, c are the weights corresponding to each factor, respectively, summing up to 1.

C. Messages Risk Assessment Mechanism

Since different users have different levels of privacy awareness, the number of sensitive attributes included in the interactive messages will be different. Information for users with the same sensitive attributes may be comfortably shared with friends in reality, but it may not be shared with strangers due to privacy concerns. Therefore, in this section, we will analyze

the messages risk rating mechanism based on user's credibility. Here we first define the sensitive attribute parameter set (SAPS) between users, as defined in Definition 3.

Definition 3 (SAPS): In all messages that a user sends to another one, they all have several sensitive attributes. The parameters of sensitive attribute i can be denoted as (α_i, β_i), where α_i denotes the attribute discrimination, and β_i denotes the attribute sensitivity. A set consisting of these sensitive attribute parameters is called SAPS.

Each user maintains an SAPS corresponding to each other user, and the determination of the sensitive attribute parameters will be discussed in the subsequent subsections.

We will rely on the method for evaluating a user's privacy score developed by LIU et al.[3], and make use of the project response theory to evaluate the message risk level.

1) Project Response Theory (IRT): Project Response Theory is a theory on psychological measurements. It is used for the project screening and test preparation. In this paper, we use an IRT model with two parameters. In this model, the characteristics of a question q_i to be measured is represented by a pair of parameters $\xi_i = (\alpha_i, \beta_i)$, where α_i denotes the distinction degree of question i, and β_i denotes the difficulty of question i. The characteristics of each examinee j are represented by a parameter θ_j, where θ_j denotes the ability of the examinee j. The basic random variable in the IRT model is to capture the quality of the examinee's answer to the question q_i. In case that the result is only "correct" or "wrong", the probability of the examinee's answer to question q_i is given in equation (5):

$$P_{ij} = \frac{1}{1 + e^{-\alpha_i(\theta_j - \beta_i)}} \tag{5}$$

This model requires three basic assumptions, namely: (1) independence between questions.(2) independence between examinees; and (3) independence between question and examinee.

2) Risk Levels of Messages: Our scheme uses the IRT model to determine the risk levels of the messages sent by users. First of all, We need to map the IRT model to suit our problem. Thus, as shown in Fig. 2, in our IRT model, we treat message receiver B as examinee j, the sensitive attribute ξ_i in the SAPS is regarded as the question q_i, and and the credibility of the message receiver is equivalent to the examinee's ability θ_i. The lower the credibility, the less sensitive the attributes. α_i and β_i correspond to the parameters of θ_i in SAPS, and P_{ij} denotes the success rate of the message receiver's acquisition of the sensitive attribute i.

In order to determine the risk level of messages sent by users, it is necessary to determine the sensitivity of all attributes and the success rate of the message receiver's acquisition of each sensitive attribute. Therefore, we need to collect all sensitive attribute parameters $\xi_i = (\alpha_i, \beta_i)$ in all messages sent by one user.

Our scheme also uses the maximum likelihood estimation method to estimate the sensitive attribute parameters $\xi_i = (\alpha_i, \beta_i)$, as shown in algorithm 1. Therefore, we need to find the maximum likelihood estimate of ξ_i to maximize the likelihood function(6),

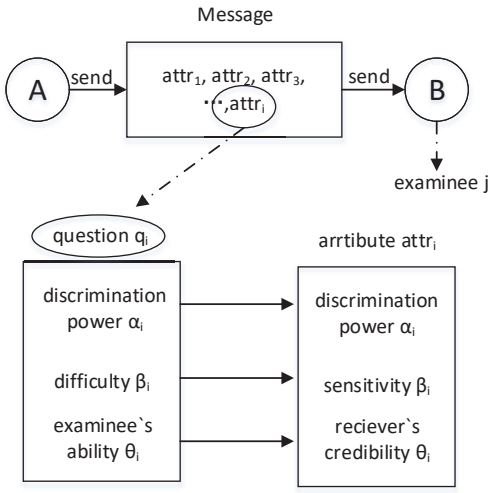

Fig. 2. Mapping relations of our scheme and IRT model

Algorithm 1 Estimation of $\xi_i = (\alpha_i, \beta_i)$ for all sensitive attributes

Input:

Messages msg, reciever's credibility $\vec{\theta} = (\theta_1, ..., \theta_N)$ and the number F of reciever's credibility groups.

Output:

$\vec{\alpha} = (\alpha_1, ..., \alpha_N)$ and $\vec{\beta} = (\beta_1, ..., \beta_N)$

1: $\{G_g, \theta_g\}_{g=1}^{F} \longleftarrow Patitionreciever(\vec{\theta}, F)$
2: for g=1 to F do
3: $\quad p_g \longleftarrow |G_g|$
4: for i=1 to n do
5: $\quad w_{ig} = |\{j | j \in G_g \bigcap Info_{set}(i,j) = 1\}|$
6: for i=1 to n do
7: $\quad (\alpha_i, \beta_i) \longleftarrow NR\_Estimation(msg, \{P_g, w_{ig}, \theta_g\}_{g=1}^{K})$

$$\prod_{j=1}^{N} P_{ij}^{Info(i,j)} (1 - P_{ij})^{1 - Info(i,j)} \tag{6}$$

where $Info(i,j)$ indicates whether the user j wants to get the attribute i of the messages. The likelihood function assumes that the credibility levels of all users are different. However, in reality, OSN users $\{1, ..., N\}$ can be divided into F non-overlapping groups $\{G_1, ..., G_F\}$, where $\bigcup_{g=1}^{F} G_g = \{1, ..., N\}$, users of each group have same credibility. Let θ_g represent the credibility of user group G_g, and p_g denotes the number of users in group G_g. For each sensitive attribute i, let w_{ig} represent the number of people who have succeeded in obtaining the other user's attribute i in the user group G_g, that is, $w_{ig} = |\{j | j \in G_g \bigcap Info_{set}(i,j) = 1\}|$. Therefore, the above likelihood function can be further expressed as in (7).

$$\prod_{g=1}^{N} [P_i(\theta_g)]^{w_{ig}} [1 - P_i(\theta_g)]^{p_g - w_{ig}} \tag{7}$$

and the corresponding log-likelihood function is shown in (8).

$$L = \sum_{g=1}^{K} (w_{ig} log P_i(\theta_g) + (p_g - w_{ig}) log(1 - P_i(\theta_g))) \quad (8)$$

We use the Newton-Raphson algorithm in maximizing the log-likelihood function, and the algorithm gives the estimated value $\xi_i = (\alpha_i, \beta_i)$ through calculating the first order derivative and the second order derivative with respect to α_i and β_i iteratively, and then, we can obtain the $\widehat{\alpha_i}$ and $\widehat{\beta_i}$.

Finally, We take user credibility θ_i and the sensitive attribute parameters ξ_i into formula(5), and use the equation (9) to calculate the risk level PR(j) of SAPS sent by user.

$$PR(j) = \sum_{i=1}^{n} \beta_i \times \frac{1}{P_{ij}} \quad (9)$$

Before the initial contact session between user A and user B, the SAPS of both sides are empty. When user A sends each message to user B, the risk level of each message is determined by using the above method. As long as user A sends a message, the system records sensitive attribute parameters contained in this message into the SAPS. With the progress of the session, the SAPS of user A will contain multiple sensitive attribute parameters gradually. When the risk level of the messages sent by user A is too high, the system will issue a warning to user A, and user A can choose whether continue to send more message to user B based on the nature of the warning.

IV. Experiment Study

Since user credibility has significant impact on the assessment of messages risk level, we will use a Facebook-like Social Network data set to evaluate user credibility in this study. Here, we simulated a number of malicious users according to the behavioral characteristics of malicious users in OSNs, and added them to the data set. More specificlly, the data set contains specific timestamps of interactive messages of 1100 users (including 100 malicious users). During the analysis, we set T to be 12 hours. Our goal is to identify as many malicious users as possible, and reduce the error recognition rate.

A. Distribution of user credibility

We can get the distribution of user credibility by calculating, as can be seen in the Fig. 3.

The number of ordinary users whose credibility in 0.1∼0.2 and 0.2∼0.3, respectively, are small, and the number of ordinary users whose credibility in 0∼0.1, 0.3∼0.4 and 0.4∼0.5 respectively, are large, and the number of ordinary users whose credibility in 0.7∼0.9 is the smallest. Fortunately, most of malicious user credibility less than 0.2, and a small part of them between 0.2∼0.4, so they are divided into low-credibility users.

After the careful analysis, we conclude that the user whose credibility higher than 0.5 have a lot of active friends, so they

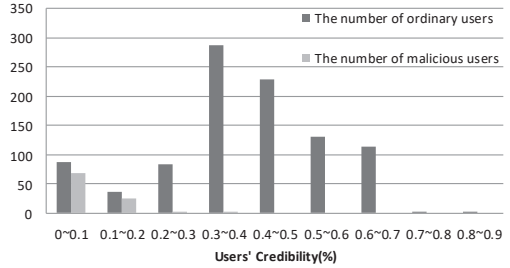

Fig. 3. Distribution of ordinary user credibility

TABLE I
DISTRIBUTION OF ORDINARY USERS' FRIENDS AND CREDIBILITY

User(No.)	Positive user (quantity)	Negative user (quantity)	User's Credibility(%)
569	2	14	0.074
36	6	75	0.071
189	3	25	0.161
204	8	58	0.179
144	19	35	0.267
32	52	86	0.372
1253	11	20	0.364
1339	16	15	0.435
372	67	37	0.460
1285	21	11	0.510

are less likely to be malicious users. Users whose credibility in 0.2∼0.5 also have many friends, but the number of active friends less than high-credibility users. Users whose credibility in 0∼0.2 are belong to low-credibility users. Although most of them have a lot of friends, but both sides do not have information interactions, and it often happens that a user sends a message and the other user does not respond.

Next, we discuss the initialization of user credibility for newly registered users. When the credibility is initialized to 0, it is difficult for newly registered users to get socially active at the very start, and so we use the average credibility of all active users in OSNs to initialize user credibility for newly registered users.

B. Ordinary User Credibility

We take No.103 ordinary user as an example, because he has a large number of friends relative to other ordinary users. He has 216 friends in total, including 53 active friends and 163 negative friends, as shown in Fig. 4(a). Although he has a lot of friends, his credibility is only 0.28. At the same time, we also analyze 10 different ordinary users randomly selected in different credibility intervals, as shown in Table I. The results show that user credibility is related to the proportion of positive and negative friends. With the increase of the proportion of positive friends, the user credibility is also improved.

C. Malicious User Credibility

We take No.2064 malicious user as an example in this analysis, because he has a great number of friends relative to

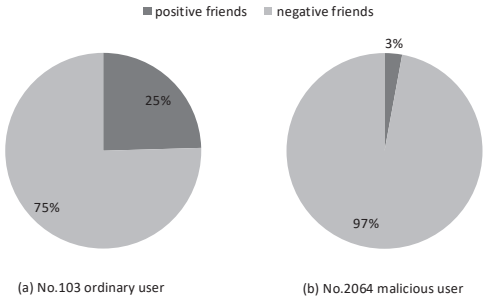

Fig. 4. The propotion of positive friends and negative friends

TABLE II
DISTRIBUTION OF MALICIOUS USERS' FRIENDS AND CREDIBILITY

User(No.)	Positive user (quantity)	Negative user (quantity)	User's Credibility(%)
2016	5	210	0.014
2038	6	258	0.025
2014	15	199	0.063
2098	16	132	0.078
2030	12	41	0.088
2048	5	183	0.093
2069	7	59	0.102
2022	9	178	0.135
2073	18	35	0.219
2089	13	28	0.323

other malicious users. He has 415 friends in total, including 12 active friends and 403 negative friends, as shown in Fig. 4b. We also randomly select 10 other malicious users for our analysis, as shown in table II. The results show that the credibility of malicious users is generally low.

With the improvement of privacy awareness, most ordinary users will disregard messages which are sent by such malicious users. Even if there are a small part of users interact with them, the frequency of interaction is low. This will lead to such malicious users have fewer positive friends and more negative friends.

V. CONCLUSION

In this paper, we propose an interactive message privacy protection scheme based on user credibility in OSNs. We measure user credibility by social behaviors, and then apply the IRT model to assess the risk level of messages. Users can not only choose whether to add friends by user credibility, but also choose whether to send messages with sensitive attributes according to other user credibility. Our scheme provides users with useful recommendation for message risk level before users send out messages that contain sensitive attributes. Through extensive analysis and experimental verification, we demonstrate that the most of malicious users are rated as low credibility users. Therefore, our scheme can effectively reduce the risk of privacy leakage in future interactive messaging systems.

ACKNOWLEDGMENT

The authors would like to thank Prof. Yuguang Fang for providing many useful suggestions for our scheme.

The work is supported by the National Natural Science Foundation of China under Grant No.61672106 and the General Program of Science and Technology Development Project of Beijing Municipal Education Commission under Grant No.KM201611232013.

REFERENCES

[1] N. B. Ellison *et al.*, "Social network sites: Definition, history, and scholarship," *Journal of Computer-Mediated Communication*, vol. 13, no. 1, pp. 210–230, 2007.

[2] Y. Wang and R. K. Nepali, "Privacy threat modeling framework for online social networks," in *Collaboration Technologies and Systems (CTS), 2015 International Conference on*. IEEE, 2015, pp. 358–363.

[3] K. Liu and E. Terzi, "A framework for computing the privacy scores of users in online social networks," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 5, no. 1, p. 6, 2010.

[4] L. Li, T. Sun, and T. Li, "Personal social screen–a dynamic privacy assignment system for social sharing in complex social object networks," in *IEEE Third International Conference on Privacy, Security, Risk and Trust*, 2011, pp. 1403–1408.

[5] A. Srivastava and G. Geethakumari, "Measuring privacy leaks in online social networks," in *Advances in Computing, Communications and Informatics (ICACCI), 2013 International Conference on*. IEEE, 2013, pp. 2095–2100.

[6] B. S. Vidyalakshmi, R. K. Wong, and C. H. Chi, "Privacy scoring of social network users as a service," in *IEEE International Conference on Services Computing*, 2015, pp. 218–225.

[7] J. Qian, X. Y. Li, C. Zhang, and L. Chen, "De-anonymizing social networks and inferring private attributes using knowledge graphs," in *IEEE INFOCOM 2016 - the IEEE International Conference on Computer Communications*, 2016.

[8] G. Petkos, S. Papadopoulos, and Y. Kompatsiaris, "Pscore: A framework for enhancing privacy awareness in online social networks," in *Availability, Reliability and Security (ARES), 2015 10th International Conference on*. IEEE, 2015, pp. 592–600.

[9] M. Otsuki and N. Sonehara, "Estimating the value of personal information with sns utility," in *Eighth International Conference on Availability, Reliability and Security*, 2013, pp. 512–516.

[10] M. Li and A. Bonti, "T-osn: A trust evaluation model in online social networks," in *Ifip International Conference on Embedded and Ubiquitous Computing*, 2011, pp. 469–473.

[11] H. Liu and M. Abbasi, "Measuring user credibility in social media," *Social Computing, Behavioral-Cultural Modeling and Prediction*, pp. 441–448.

[12] L. Guo, Y. Fang, and L. Wei, "Fine-grained privacy-preserving reputation system for online social networks," in *Communications in China (ICCC), 2013 IEEE/CIC International Conference on*. IEEE, 2013, pp. 230–235.

[13] M. Gambhir, M. Doja, *et al.*, "Action-based trust computation algorithm for online social network," in *Advanced Computing & Communication Technologies (ACCT), 2014 Fourth International Conference on*. IEEE, 2014, pp. 451–458.

[14] T. Reynaert, W. De Groef, D. Devriese, L. Desmet, and F. Piessens, "Pesap: A privacy enhanced social application platform," in *Ase/ieee International Conference on Social Computing and 2012 Ase/ieee International Conference on Privacy, Security, Risk and Trust*, 2012, pp. 827–833.

[15] E. Palomar, L. González-Manzano, A. Alcaide, and Á. Galán, "Implementing a privacy-enhanced attribute-based credential system for online social networks with co-ownership management," *IET Information Security*, vol. 10, no. 2, pp. 60–68, 2016.

[16] C. Li, B. Palanisamy, and J. Joshi, "Socialmix: Supporting privacy-aware trusted social networking services," in *Web Services (ICWS), 2016 IEEE International Conference on*. IEEE, 2016, pp. 115–122.

[17] K. W. Ahmed, I. J. Mouri, R. Zaman, and N. Yeasmin, "A privacy preserving personalized group recommendation framework," in *Advanced Computing (IACC), 2016 IEEE 6th International Conference on*. IEEE, 2016, pp. 594–598.