A Network Registration Model with Hidden Real Name

Kan Chen
College of Computer
National University of Defense Technology
Changsha, China
jeffee@nudt.edu.cn

Peidong Zhu
College of Computer
National University of Defense Technology
Changsha, China
pdzhu@nudt.edu.cn

Yueshan Xiong
College of Computer
National University of Defense Technology
Changsha, China
ysxiong@nudt.edu.cn

Abstract—**The openness and anonymity of Internet allow people joining in without identity validation, resulting in a trustless environment and bringing about many security problems, such as privacy exposure, rumor spreading, deception and social panic making. The root of these problems is that users use pseudonyms instead of their real identities, thus they take no responsibility to their misbehaviors. To solve the problems and build a more credible environment, real name policy is promoted in some countries. Although the real name policy can provide strong assurance to users' real identities, but as a compelled governmental law, it suffers difficulty in promotion and even worse, may lead to the disaster of large-scale leaking of real name information. In this paper, we provide a novel real name validation model. The key of our model is that users are validated by their friends and their real name information is kept by their friends. In the way, the server is free to the burden of storing privacy information and the real name is hidden to the server and other users. We test our model with real data and get inspiring results.**

Keywords-real name; network registration; model; relationship

I. INTRODUCTION

Due to its open and anonymous design, the Internet provides enough freedom and convenience for public to use. Everyone can join in a network after registration with a nickname, which is used as a description of his network identity, but is unrelated with his real identity. Thus, the identities in network are always trustless, resulting in a great many of problems.

In a network setting, the identity problems can be mainly summarized as follows.

- You don't know who you're talking about.
- You don't know whether you're talking with the one you want.
- You don't know who should be responsible for a networking talk.

These problems can also result in some worse consequences. For example, malicious users pretend to be friends for deception. It's hard to distinguish because of the lack of verify mechanisms. Besides, the anonymity of Internet provides a platform to speech without taking responsibility, which brings about more baseless rumors, personal defamation, and privacy invasion than in real life.

In order to conquer the problems and build a trust environment, some countries refer to the real name policy. Under this compelled law, users are forced to register with their real name when they come into some networks. The key point of real name policy is that users are traceable. Once misbehaving, users can be located and supervised according to his registered information.

Real name policy can prevent some criminals from illegal behaviors, such as spreading fraud information, creating social panic, phishing and deception. But it also has fatal shortcomings. As the real name information is kept in server, it's easy to be targeted by malicious attackers and face the threat of large-scaled real name information leaking. In fact, as the first country using real name policy, the South Korea suffered such disasters in 2006.

In this paper we provide a novel registration model. The contribution of our work is three folds. First, the server doesn't need to store all the users' real name information, which extremely reduces the threat of privacy leaking. Second, users do not need to provide their real name information to the server, which eliminates their concerns about showing real name in network. Third, although the server does not store the real name of the users, it can trace the users' real identity when necessary, such as when the user publishes some illegal contents.

In a word, our model leverages the trust relationships in online social network to construct verification links. Users' real name information is validated by their trustful friends. In this way, users' privacy is hidden to the server and the others, which reduces the threat of privacy invasion.

The remainder of this paper is organized as follows. The related works are introduced in section 2. We describe our model in detail in section 3. The experiment is shown in section 4 and the discussion is in section 5.

II. RELATED WORKS

Internet is an open and free platform. Many services allow anonymous access[1]. Some others ask users to register at the first visit and then access with usernames. Even so, most of them do not care about users' real identities. As a result, almost all the users register with pseudonyms, which have nothing to do with their real identities[2].

Using pseudonym is a kind of protection to users' privacy[3]. But it also provides convenience to commit criminal stuffs. A direct consequence of anonymity and pseudonymity is that people can publish anything in Internet without taking responsibilities, because their real identities are untraceable. As a result, many incidents have been introduced, such as deception, defamation, slander and social panic making.

In order to solve the referring problems, the South Korea government published the Real Name Verification Law in July 2007. According to the law, users should validate their real names to register in a websites. It's the beginning of public use of real name policy[4].

The original policy goal is to prevent widespread online abuse in postings and comments that may cause serious privacy invasion and personal defamation. It's found that it indeed shows significant effects on reducing misbehaviors[5]. But the side-effect is fatal. As the real name information is stored in the server, it's easy to be targeted by attackers and cause large scaled information leaking. In fact, since 2006, many real name leaking incidents have been reported in South Korea. In order to reduce the threat of privacy exposure, an alternative identity named i-Pin ID is then used to validate user's identity[6]. But it's still not safe against phishing attacks[7].

China has been attempting to promote real name policy for years. Presently it has been launched in the area of e-commerce, transportation, online games and some online social networks[8, 9].

Many designers believe that the real name policy is necessary to encourage healthy interactions in online communities[10, 11]. But it's not embrace by all[12, 13]. The doubt mostly originates from the threat of privacy leaking. There've been some research on extra security mechanism[14], but as long as the real name information is kept by the server, the threat of privacy invasion is potential and inevitable.

III. SYSTEM MODEL

The main concern of real name policy is that it may give convenience to attackers for information stealing. The accidents happened in South Korea have proven that the threat indeed exists. But if we free the server's burden of storing sensitive information, the threat would be eliminated.

The key point of our model is that user's real name information is no longer stored by the server. Instead, it's validated and maintained by their trusty friends. Thus, the privacy is hidden to the server and the others but the users' identities can still be validated and traced.

A. User Category

In China, real name policy has been tested in some specific areas for several years. Many users have registered with their real names, but there're still some with doubtful attitudes. And the concerns mostly result from distrust to the Internet environment.

It has been found that trust is strongly related to information disclosure[15]. In real life, people are more likely to trust each other than in network space. The reason is that in real life, people contact with each other directly, they get more information from each other. If they become friends, they trust each other more and share more information such as name, age, address, email, phone number et al for communication and politeness.

This is the basic of our work. We assume that people prefer to trust their trusty friends and are willing to give their real names to their friends. It's reasonable in real life. If a person does so, we say that his identity is verified by his friends, and his friends are his verifiers. Here we have a restriction to the verifier. If someone wants to be verifier of another, he should be validated first.

In our model, we combine server validation and friends validation together. It's because someone may want to be validated with real name. For example, some celebrities want to use their real name to communicate with fans in Internet.

According to the type of validation, users are categorized as the authorized, the verified and the unverified.

1) Authorized User: An authorized user is registered with his real name. Server keeps his personal information confidently. The authorized users have the most credible identities.

2) Verified User: A verified user is validated by the authorized or other verified users. Their identities are less credible than the authorized.

3) Unverified User: An unverified user is not validated by anyone authorized users or verified users. Thus his identity is not credible.

The authorized users and the verified users are all validated. The difference lies that their verifiers are different and their real name information are kept differently.

B. User Validation

When registering in online social networks, user can choose whether to register with his real name or not. Nowadays, many networks encourage users to register with real names because it's more convenient for the networks to provide personal services, such as finding old friends and recommending new friends.

People do not want to register with real name for several reasons. First, they do not want to show their real identities in Internet. Second, they worry about information leaking. If they still want to be validated, they can choose to be verified by their validated friends.

Let's represent *U* as a new user, and *S* as the server. If *U* wants to be validated by *S*, he follows the procedures below.

1) Request Asking: *U* send validation request to *S*;

2) Candidates Feedback: *S* checks all of *U*'s friends and picks up those who have already been validated into list *L*, and then feeds back with *L* to *U;*

3) Verifiers Choosing: *U* chooses at least m friends in *L* as his verifiers and add them to set *VS(U)*, then sends request to each of *VS(U)* with information about *U*'s real name;

4) Identity Assurance: Every member of *VS(U)* check the content of *U*'s request and judge that whether the information is true according to their knowledge about *U*. If the answer is positive, a verification message is fed back to *S*. Otherwise, the verifier just ignore the request and does not give any feedback;

5) Coherence Checking: Once *S* has received more than *m* positive feedbacks, it notices *VS(U)* to check the coherences of their knowledge about U. If all the verifiers get the same information, they give positive feedback to *S*, and then *S* stores *U* and *VS(U)* in its database. By then, *U* has been successfully verified. If *VS(U)* find that they get different information from *U*, they will discard the current procedures and ask *U* to resend new requests.

If *U* succeeds in validation, his real name is maintained by his verifiers, and the relationship information is kept in the server.

In this way, users are validated by their trusty friends and their real name information is kept by their friends. It's safe because user only choose those they believe as their verifiers. Besides, it's more flexible and accurate because after registration, if the real name information changes, such as house moving or phone number changing, few would update their registered information but most would like to notice their friends.

C. User Trace

The benefit of real name lies that users are traceable after registration with real name. If someone publishes illegal content in network, he can be traced by his registration information and supervised by relevant regulations.

In our model, only the authorized users' real name information is kept in server. On the contrary, the verified users only store information about the relationship with their verifiers in server. In using of these relationship information, we can still get the real name of a verified user.

Consider a topology as Figure 1. The arrows indicate the direction of validation. For example, there's an arrow from U_1 to U_3, which means that U_3 was validated by $U_{1,}$ and U_1 is a verifier of U_3.

Let's assume that U_7 has misbehaved and some related regulation, for example the police, wants to get the real name of U_7 for inquisition. We'll show how this achieves.

First, we check the database of server to find that U_7 was validated by an authorized user U_6 and a verified user U_5. Once we get one of them, we can get the real name information of U_7. Here we want to get it from both U_5 and U_6 in case that one of them may tell untruthfully.

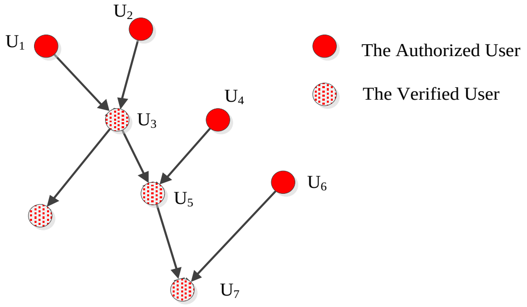

Figure 1. topology of a network

U_6 is an authorized user with information in the server. We can get it directly. Then we refer to U_3 and U_4 for U_5's information because they're U_5's verifiers. We find that U_4 is authorized and U_3 is verified. Similarly we turn to U_1 and U_2 for U_3's information. As U_1 and U_2 are all authorized, we can get their information directly. So now we can contact with U_1 and U_2 to get the Information of U_3. The messages from U_1 and U_2 are checked and compared to get U_3's real name information. After contacting with U_3, in the same manner, we can get U_5's information and finally get U_7's real name.

Here it's worth to say that every verifier should ensure that his warrantee has provided true information when asking for validation. But he takes no responsibilities for his warrantee's behaviors. That's to say, the misbehaviors of a user wont' affect his verifier. This is important because someone may not be willing to be a verifier because of worrying about being entangled.

IV. Experiment

To validate the effectiveness of our model, we experiment with real data from Weibo, the leading micro-blog service of China. Weibo encourage users to validate with real names. Those who have been validated are marked with a symbol "V". We can get that whether a user is validated or not through the descriptions in his homepage.

TABLE I. Statistical Details of G

Name	Value	Description
#(V)	13,420	Number of users
#(E)	209,857	Number of relations
#(Verified)	372	Number of Verified users

We crawled 13,420 users and their relationships from Weibo. The users and the relationships can be described as graph *G={E,V}*, while *V* represents the users and *E* refers to the relations. The statistical details of G are listed in table 1.

In our dataset, we get 372 validated users and the initial validation rate is 2.8%. We can see that most of the users did not validate their real names yet.

We refer the validated users as the authorized in our model. The rest are unverified users. For an unverified user, he is capable to be validated if he has more than *m* validated friends. We simulate the procedure of validation iteratively and calculate the validation rate. Here, we take *m*=4.

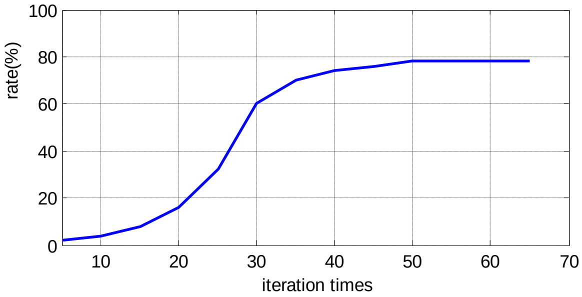

Figure 2. Distribution of validation rate from iteration times

Figure 2 shows distribution of the validation rate from the iteration times. We can see that at beginning only less than 3% users are registered. After iteration for about 50 times, nearly 80% users have been validated. It suggests that our model can validate most of the users in real network.

However, we can still find that there're 20% users are not validated. The reason is that in online social network, there're always some lazy accounts, which means that the users owning the accounts seldom access and maintain. Thus they have poor relations with others and can't satisfy the conditions to be validated. If they want to be verified, they should get better relationship with others.

From the experiment we can see that our model can validate most of the users in real networks. Although there're still some users can't be validated, the reason lies their poor relations with others. If they want to be validated, they can choose the server way or pay more attention on communication with others to get better relations. From this aspect, our model can encourage users to be more active in contacting.

V. Discussion

In this paper we introduce a novel network registration model. Our model can provide nearly equal effectiveness to the real name policy, but it's safer because the real name information is hidden to the server and the others. The experiment shows that our model can work well in real network. It can be used with the real name policy together and is easy to be adopted by the present networks.

Acknowledgment

This work is supported by the National Natural Science Foundation of China under Grant No. 61170285.

References

[1] A. Pfitzmann and M. Koehntopp, Eds., Anonymity, unobservability, and pseudonymity—a proposal for terminology. Lecture Notes in Computer Science Volume 2009, 2001, pp 1-9.

[2] I. Goldberg, Ed., A pseudonymous communications infrastructure for the internet. UC Berkeley Working Paper, 2000.

[3] A. Teich, et al., "Anonymous communication policies for the Internet: Results and recommendations of the AAAS conference," The Information Society, vol. 15, 1999.

[4] J. Leitner, "Identifying the Problem: Korea's Initial Experience with Mandatory Real Name Verification on Internet Portals," J. KOREAN L, vol. 83, 2009.

[5] D. Cho, et al., "Empirical Analysis of Online Anonymity and User Behaviors: The Impact of Real Name Policy," in System Science (HICSS), 2012 45th Hawaii International Conference on, 2012, pp. 3041-3050.

[6] i-pin. Available: http://i-pin.kisa.or.kr/

[7] Y. Oh, et al., "Empirical analysis of Internet identity misuse: case study of South Korean real name system," in Proceedings of the 6th ACM workshop on Digital identity management, 2010, pp. 27-34.

[8] Fong, S.; Yan Zhuang; Lu, L.; Rui Tang, "Analysis of general opinions about Sina Weibo micro-blog real-name verification system," Future Generation Communication Technology (FGCT), 2012 International Conference on , vol., no., pp.83,88, 12-14 Dec. 2012

[9] King-wa Fu; Chung-hong Chan; Chau, M., "Assessing Censorship on Microblogs in China: Discriminatory Keyword Analysis and the Real-Name Registration Policy," Internet Computing, IEEE , vol.17, no.3, pp.42,50, May-June 2013

[10] D. Boyd, "The politics of real names," Communications of the ACM, vol. 55, pp. 29-31, 2012.

[11] David Davenport. 2002. Anonymity on the Internet: why the price may be too high. Commun. ACM 45, 4 (April 2002), 33-35.

[12] D. Cho, "Real Name Verification Law on the Internet: A Poison or Cure for Privacy?," Economics of Information Security and Privacy III, pp. 239-261, 2011.

[13] A. Teich, et al., "Anonymous Communication Policies for the Internet: Results and Recommendations of the AAAS Conference," The Information Society, vol. 15, pp. 71-77, 1999.

[14] W. Hu and J. Chang, "Design and Implementation of the Internet Real-Name Authentication System Based on Public Key Infrastructure," in Management and Service Science (MASS), 2010 International Conference on, 2010, pp. 1-4.

[15] M. J. Metzger, "Privacy, Trust, and Disclosure: Exploring Barriers to Electronic Commerce," Journal of Computer-Mediated Communication vol. 9, 2004.