Social Network Privacy Protection Research in Big Data Era

Shuang LIANG
Shenyang Institute of Technology, Information and Control College, Liaoning Fushun China
ls_happiness@163.com

Keywords: Big Data, Social Network, Privacy Protection, Information Security.

Abstract. With the advent of the big data era, people's lifestyles are quietly changing. Social networking has become an Indispensable part, but there is a great leaks risk of user's privacy in social network. Based on the problems of big data, social network big data and privacy data on the social network, this paper has analyzed privacy protection methods and problems in traditional social network, described the current privacy protection research status and its key technologies of social network in current big data environment, given the future development direction. Only the effective privacy protection technologies combined with related policies and regulations, the problem of personal user information security can be solved. It can provide protection for the healthy development of social networks on big data era.

Big Data and Privacy Protection in Social Networks

Big Data of Social Network

As the development of wireless communication technology and popularity of 3G, 4G users, mobile Internet has emerged. social networks platforms have a new direction mobile social network. Global social network, such as Facebook, Twitter, Linkedin, QQ, WeChat, has formed more than one billion people in the cross-border, cross-ethnic and cross-cultural community of new network [1].

Mobile social network privacy in big data era

Data security and privacy issues is one of the recognized key technical problems in big data era[2-3]. Currently, any user action on the network are vested in the hands of businesses, including location track, research habits, eating habits, shopping habits, reading habits, friends contacted habits ect.. These data that was analyzed can expose personal privacy [4]. If businesses published these data, personal sensitive information will be exposure, personal privacy will be infracted. If these data cannot be published, it cannot realize to share data. So how we do not only publish more data but also cannot leak personal privacy, ensure the balance of data opening and privacy protection, it is the problem we need study.

Privacy Protection Method and Problem in Traditional Social Networks

User privacy protection usually finishes by setting privacy rights their own in traditional social network. For example, it can set which users can view the log, album, talk, sharing and video. Each piece of information may be saw for every person, only QQ friends or only by own. You can set verify way when you add friends. You can choose to let everyone to be friend, can choose verify information, can choose answer question correctly, can choose answer question and audit, can choose not allowed any person as a friend.
The above safety setting looks like covered all possible users want to control. If everyone can set carefully their own privacy authority, the information can be seen by needed person. However, most users don’t know these setting, even some don’t know these function, thus put their privacy at a very low level configuration, so that almost all of the network users can gain them. Certainly, these have harmed personal privacy seriously. Even if users can set their authority to let the information can be seen by who just wanted to let people see what he saw. But all information are in service provider, the provider can still use this information to do other things.

**The Typical Social Network Security Technology**

**Location Privacy Protection Technology**

The paper [5] focused on the user's location and its relationship to publish content in social networks, present new security solutions for location-based services brought about privacy disclosure.

If attacker has grasped user’s historical position, and calculated position data leakage risk. The judge standard whether data publisher is leakage in present position, it can be taken positions big data protection technology based on the probability speculative [6]. It mainly used Markov model [7] and generalized model graph model [8] to describe the user location data relationship between in space and at time.

If user has higher level privacy requirements, it can take location big data privacy protection technologies [9] based on privacy information retrieval. Its basic idea is users calculated locally, the data which is needed get from server by PIR [10-11]. Server without knowing the user's location data are available to give query results.

In the big data environments, an attacker can obtain user data from multiple sources, not just the same source release, this will be the direction of future research about the position privacy protection.

**Identifier Anonymity Protection Technology**

Mobile social network installs third-party software on your mobile device. Mobile users access the corresponding user service through third-party servers. Because the transmission of data uses a wireless network which is opened and user identity can easily be stolen. Regarding the issue above, it has proposed the anonymous ID and one failure in network transmission in paper [12-14], this method can ensure that even if an attacker intercepts an anonymous ID, it can’t attack social network. Aimed at the problem which once the attacker intercepts a part of user information, it is possible to analyze the corresponding one or more user's problem, the paper [12] designed a user privacy protection scheme based on the model with K-Anonymity [14]. The scheme has determined the minimum user information combination number by setting privacy strength coefficient.

It has proposed the corresponding solutions in paper [15-16] adjective node as long as gets user’s part information can do some attack after the massive data release.

**Connection Anonymity Protection Technology**

For AOL’s data leakage problem, it has proposed no center social networks based on P2P to prevent the massive data leakage from center server in paper[17-19]. Every node is as a user in this
paper. The anonymous ID or node identifier can be identified uniquely, and take friend nodes as its neighbor nodes in improved DHT algorithm [20]. Information storage, transfer and finding take public infrastructure as a security guarantee. It is not only ensures the anonymity, but also improves the search efficiency.

There is not only centralized data leakage, but also a "asymmetric settings" data leakage in the mobile social network. The so-called "asymmetric settings" is user A has set its private information, it permits part user B can access A’s private information. But user B can’t set private safe, leading to user information through the user a B leaked. For such problems, it has proposed symmetric privacy settings scheme in paper [20]. It has taken weak link to predict possible network for relationship rarefaction in paper [21]. It has predicted for hierarchical relationships in existing group in paper [22]. It has predicted and analyzed for blog social network in paper [23]. It has predicted connection relationship probability by limited random walk method in paper [24].

The tightness of connection relationship in social network has important impact on the accuracy of relationship predict. The more tighten connection can lead to the more accurate of relationship predict. It should be more effective against such attacks with the increasing of data tightness in big data era.

Privacy Protection Technology Based on Data Access Control Technology

There are similar for privacy protection schemas in paper [25] and [27]. They have already provided access control technologies based on group or role. The difference is that, the schema in paper [25] combined with attribute based encryption, ABE [26] and the traditional public key cryptography. It allows user to store the personal information to, and the third parties are not necessarily to be trusted by the user, but it does not provide group revoked schema. It has proposed role schema [27]. The schema is a access control scheme for the condition of user privacy information stored in third party, and using public key encryption theory such as broadcast encryption, keyword search ect.. At the same time it has provided convenient establishment and revocation solution for dynamic group membership changes, and improved the search efficiency.

Privacy protection technology based on data access control often needs to define and quantify data risk in a particular environment, but in big data environments, this definition and quantification will be increasingly difficult.

Privacy Protection in Social Network and its Development and Countermeasures Analysis in Big Data Environments

Privacy Protection in Social Network and its Development in Big Data Environments

With social networks in people's daily lives and more widely applied, the privacy protection of social networking has become the focus of researchers studied. The current research is mostly privacy protection technology based on information publishing such as location privacy protection technology, identifiers anonymous protection technology, connection relationship anonymity protection technology which are information publication privacy protection technology. Current technology is dependent mostly static data, but in the real world, data
patterns, and content are changing constantly. Therefore, dynamic data privacy protection will be the direction of future privacy protection in the condition of more amount data and more complex environment.

Depending on the application scene, based on the group attribute and role access, taking advantage of user's group membership, exploring information sharing program for user-group management and user-defined access policy, so that the information can be more convenient sharing within the group, in order to ensure that the information can only be seen by what people want.

Characteristics of social network are users’ relationship is complex, and the amount of data is huge. Characteristics of cloud computing are that can process a huge amount of data. Therefore, both sides has combined effectively, to design a more rational social network architecture based on cloud computing, and put forward to safe information sharing and storage mechanism, it will also be an important direction for future development of social network privacy protection.

The Other Countermeasures to Improve Social Online Privacy Security

Social network privacy protection in big data era, developing the research under the premise of no reducing data open, continuously improve the strength of privacy is the important aspect of protecting personal privacy. At the same time it can speed up the pace of construction in the following areas.

Personal Information Security in Law Synchronous. Big data technology is a new technology in recent years. Therefore, strengthen relevant laws are being implemented. Although it has promulgated "information security technology, business services and public information systems Personal Information Protection Guide" in March 1, 2013. It is the highest national standards as personal information protection. But it cannot meet the specifications required of information control by people needed. So the basic legal for protecting personal basic information is of great urgency.

Specification Conducting Standards of Social Network Trade. In order to get more benefits from big data in social networking industry, it is necessary to build customer trust, maintain user information security, build industry standards effort. Collecting user information in the condition of user fully aware and ensure that user information does not user for other purposes. Try their best to seek the balance among user, data service provider and data service consumer. It must be ensure legitimate and reasonable using data. It will realize data maximum share under the premise of making safe data environment.

Raise User's Privacy Protection Awareness. Raise awareness of user privacy is proactive effective measures to protect personal information in the big data era. When User applies social networking, they can share their real photo limited, personal address location, personal records of daily life etc., effectively set strangers’ access permissions, avoid disclosing an individual's name, birth-date, address, phone-number, email and other information in network accounts, read security privacy protocol of social network operator carefully, set privacy protection correctly, to proactive protect personal information security.
Conclusions

Social network privacy protection have new features in big data era. It is a contradiction that is necessary to ensure that a large number of data is shared, but also to ensure effective protection of user privacy. This paper has introduced several major technology main issues and its developing progress such as location privacy, anonymity identifier, connection relationship, identity-based privacy protection, accessing-control-based privacy protection and group- attribute-digging privacy protection. However, these studies is not enough. Through the analysis of the previous research, this paper has given some directions for future research and other policies. The author believes that combining relevant laws policies and technical means can solve social network security and privacy issues effectively.

References


