Resolving Multiparty Privacy Conflicts in Social Media

Mr. A. J. Patankar[1] Neha Chiddarwar[2] Mayuri Kulkarni[3] Chaitali Kulkarni[4] Sayali Kharche[5]
[1]Assistant Professor [2,3,4,5]BE Student
[1,2,3,4,5]Department of Information Technology
[1,2,3,4,5]DYPCOE, India

Abstract— On-line social networks like Facebook are increasingly utilized by many people. These networks allow users to publish their own details and enable them to contact their friends. Some of the information revealed inside these networks is private. These structures allow clients to present specific of them and interface with their mates. Client profile and family relationship relations are really private. These networks allow users to publish details about themselves and to connect to their friends. Some of the information revealed inside these networks is meant to be private. A privacy breach occurs when sensitive information about the user, the information that an individual wants to keep from public, is disclosed to an adversary. Private information leakage could be an important issue in some cases. And explore how to launch inference attacks using released social networking data to predict private information. In this we map this issue to a collective classification problem and propose a collective inference model. In our model, an attacker utilizes user profile and social relationships in a collective manner to predict sensitive information of related victims in a released social network dataset. To protect against such attacks, we propose a data sanitization method collectively manipulating user profile and friendship relations. The key novel idea lies that besides sanitizing friendship relations, the proposed method can take advantages of various data-manipulating methods. We show that we can easily reduce adversary's prediction accuracy on sensitive information, while resulting in less accuracy decrease on non-sensitive information towards three social network datasets. To the best of our knowledge, this is the first work that employs collective methods involving various data-manipulating methods and social relationships to protect against inference attacks in social networks.

Key words: Online Social Networks (OS Ns), Collective Inference, Data Sanitization

I. INTRODUCTION

The rapid growth and ubiquity of online social media services has given an impact to the way people interact with each other. Online social networking has become one of the most popular activities on the web. Social network analysis has been a key technique in modern sociology, geography, economics, and information science. The data generated by social media services often referred to as the social network data. In many situations, the data needs to be published and shared with others. Social networks are online applications that allow their users to connect by means of various link types. As part of their professional network; because of users specify details which are related to their professional life. These sites gather extensive personal information, social network application providers have a rare opportunity direct use of this information could be useful to advertisers for direct marketing. Publish data for others to analyze, even though it may create severe privacy threats, or they can withhold data because of privacy concerns, even though that makes the analysis impossible. A privacy breach occurs when sensitive information about the user, the information that an individual wants to keep from public, is disclosed to an adversary. For examples, business companies are analyzing the social connections in social network data to uncover customer relationship that can benefit their services and product sales. The analysis result of social network data is believed to potentially provide an alternative view of real-world phenomena due to the strong connection between the actors behind the network data and real world entities. Social-network data makes commerce much more profitable. On the other hand, the request to use the data can also come from third party applications embedded in the social media application itself. For instance, Facebook has thousands of third –party applications and the number is growing exponentially. Even though the process of data sharing in this case is implicit, the data is indeed passed over from the data owner (service provider) to different party (the application) The data given to these applications is usual not sanitized to protect users' privacy. Desired use of data and individual privacy presents an opportunity for privacy-preserving social network data mining. That is, the discovery of information and relationships from social network data without violating privacy.

Privacy concerns in social networks can be mainly categorized into two types: inherent-data privacy and latent data privacy. Inherent-data privacy is related to sensitive data contained in the data profile submitted by users in order to receive data-related services.

II. LITERATURE SURVEY

A. Inferring Privacy Information from Social Networks

Using a Bayesian network approach to model the causal relations among people in social networks. Results reveal that personal attributes can be inferred with high accuracy especially when people are connected with strong relationships.

B. You Are Who You Know: Inferring User Profiles in Online Social Networks

Using fine-grained data taken from two large online social networks, we found that users are often friends with others who share their attributes. The attributes of users, in combination with the social network graph, be used to predict the attributes of another user in the network.

III. FUTURE SCOPE

The system is about developing an online social network which provides more secure privacy of a user's profile. The system also deals with developing a mechanism to provide users with secure communication.

A. Proposed System

In this paper, we focus on latent-data privacy. We assume third party users may collect anonymous user data from social networks. Some users disclose their sensitive information, while others do not. However, third party users can carry out de-anonymization actions and further infer sensitive information of users. We first investigate how to infer sensitive information hidden in the released data. Then, we propose some effective data sanitization strategies to prevent information inference attacks. On the other hand, the sanitized data obtained by these strategies should not reduce the valuable benefit brought by the abundant data resources, so that non-sensitive information can still be inferred and utilized by third party users. To launch an inference attack by third party users, we employ a typical inference attack, called collective inference, as a case study. We present a novel implementation method for collective inference. Collective inference mainly rely on iteratively propagating current predicting results throughout a network to improve prediction accuracy, thus we need to consider how to best predict sensitive information in each iteration.

1) Advantages of Proposed System

1) Detect collective attacks in diverse large scale social networks.
2) Proposed system can work reasonably to balance privacy and data utility.
3) Third party users cannot obtain necessary information to accurately predict sensitive information.
4) Consider the special features of social network data to investigate collective attacks in diverse large scale social networks.

B. System Specification

1) Hardware Requirements

- Processor: Pentium IV 2.4 GHz.
- Hard Disk: 40 GB.
- Monitor: 15VGA

2) Colour

- Mouse: Logitech.
- Ram: 256 Mb

3) Software Requirements

- Operating system: Windows XP/7
- Coding Language: JAVA/J2EE, Hibernate.
- IDE: Java eclipse.
- Web server: Apache Tomcat 7.
- Front End: JSP, CSS etc.
- Back End: MySQL as database server.

C. System Architecture

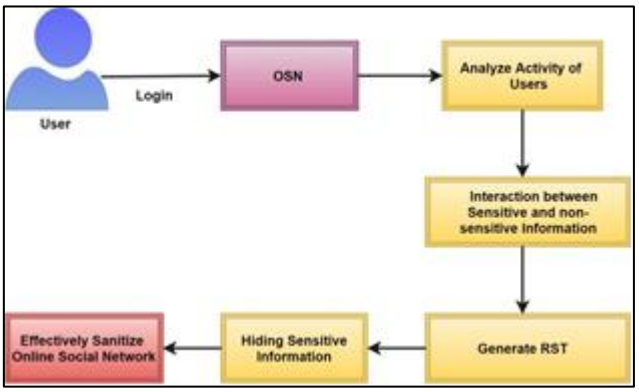

Fig. 1: System Architecture of Proposed System

D. Algorithms

1) Mathematical Model

Let W be the whole system which consists: W= {IP, PRO, OP} Where,
IP is the input of the system.
A) IP= {U, C, R, OSN, SA, UA}
U is the number of users in the system.
R is the set of number of registered U in the system.
C is the custom setting for all U.
OSN is the system.
SA is the sensitive attributes.
UA is the User Activities.

2) Naïve Bayes Algorithm Steps

1) Step 1: It is loaded the text file which will be classified as being ONE, TWO or THREE
2) Step 2: There are loaded the text file found in the folder. The name of the files belonging to class ONE are: "abc_.txt", the ones belonging to class TWO are: "pqr_.txt" and the ones for class THREE are: "xyz_.txt".
3) Step 3: It is determined the a priori probability for each class: P(UNU) = Nr Template In Class ONE / Number Total Templates P(DOI) = Nr Template In Class TWO / Number Total Templates
   P(TREI) = Nr Template In Class THREE / Number Total Templates
4) Step 4: It is determined the probability that the text file from the Step 1 to be in class ONE, TWO or THREE. Let (i,j) be the position of a attribute in the text file. It is calculated the probability that the attribute having the coordinates (i, j) to be for the class ONE, TWO and THREE.
   count1i,j = 0
   for k = 1,n ; n – the number of attribute in class ONE
   if abc_k(i,j) = 255 then
   count1i,j = count1i,j + 1
   probability1(i,j) =count1i,j / NrTemplateInClassONE
   count2i,j = 0
   for k = 1,n ; n- the number of attribute in class TWO
   if pqr_k(i,j) = 255 then
   count2i,j = count2i,j + 1
   probability2(i,j) =count2i,j / NrTemplateInClassTWO
   count3i,j = 0
   for k = 1,n ; n- the number of attribute in class THREE
   if xyz_k(i,j) = 255 then
   count3i,j = count3i,j + 1

probability 3(i,j) =count3i,j / NrTemplateInClassTHREE

5) Step 5: The posteriori probability that the attribute in Step 1 to be in class ONE is:
   P(T|ONE) = average (probabilitate1(i,j)); (i, j) – the position of the attribute in the text file from Step1
6) Step 6: The posteriori probability that the attribute in Step 1 to be in class TWO is:
   P(T|TWO) = average (probabilitate1(i,j)); (i, j) – the position of the attribute in the text file from Step1
7) Step 7: The posteriori probability that the attribute in Step 1 to be in class THREE is:
   P(T|THREE) = average (probabilitate1(i,j)); (i, j) – the position of the attribute in the text file from Step1
8) Step 8: It is determined the probability P for each text file class and it is assigned the text file from Step1 to the class of text file that has the greatest probability. P(ONE|T) = P(T| ONE)*P(ONE)
   P(TWO|T) = P(T| TWO)*P(TWO)
   P(THREE|T) = P(T| THREE)*P(THREE)

E. PRO is the Procedure of our Proposed System

1) Step 1: At first user will register into the OSN system with his/her basic information.
2) Step 2: The registered information will be forwarded to OSN system.
3) Step 3: OSN system will check the sensitive and non-sensitive attributes of registered users.
4) Step 4: OSN system will automatically hide the sensitive Information system.
5) Step 5: Then user will login into the system.
6) Step 6: User will perform the UA like profile setting, post sharing, like or comment onto the post and message sending to the another users by matching the attributes.
7) Step 7: Then OSN will provide the privacy for users likes and comments post.

OP is the output of the system:
The system provides the privacy to the user's sensitive data and privacy for posts which share by users.

1) Modules

1) User
2) OSN System
   a) User

– Registration Login
– Post Status
– Profile setting
– Send message to another
– Users Logout

F. Registration

The user will register to the system with normal information. At the time of registration the OSN system will hide the user's sensitive information.

G. Login

Fig. 2: Login

For login to the system, user will enter the Username and password, if entered details are correct then the system will redirect him to home page otherwise it will shows an error message.

Fig. 3:

1) After Login

– User will share the post.

2) Post the Status

Set the setting to profiles.
Send the messages to other users by checking the attributes.

H. Logout

User logout the account from system.

1) OSN System

The OSN system:

– Check sensitive and non-sensitive information of all users
– Check the all registered users sensitive information. It stored the sensitive attributes.
– The OSN will provide the privacy for users like and comments posts.

ACKNOWLEDGMENT

We have taken efforts in this project, however, it would not have been possible without the kind support and help of many individuals and organizations. We would like to extend our sincere thanks to all of them. We are highly indebted to Mr.A.J.Patankar for his guidance and constant supervision as well as for providing necessary information regarding the project & also for his support in completing the project. We would like to express our gratitude towards our parents & our Head of I.T. Department Dr.PreetiPatil for their kind co-operation and encouragement which helped us in completion of this project. Furthermore, I would also like to acknowledge with much appreciation the crucial role of the staff of DYPCOE Akurdi, who gave the permission to use all required equipment and the necessary materials to complete my project stage I. We are also deeply grateful to the Principal of DYPCOE, Dr.VijayWadhaiand my parents for their financial and logistical support and for providing necessary guidance concerning project's implementation.

REFERENCES

[1] j. he, w. chu, and v. liu(2006), "Inferring Privacy Information from Social Networks," Proc. Intelligence and Security Informatics.
[2] E. Zheleva And L. Getoor(2008), "Preserving The Privacy Of Sensitive Relationships In Graph Data," Proc. First AcmSigkdd Int'l Conf. Privacy, Security, And Trust In Kdd, Pp. 153-171.
[3] S. Nilizadeh, A. Kapadia, and Y.-Y.Ahn, "Community-enhanced de-anonymization of online social networks," in Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, ser. CCS '14. New York, NY, USA: ACM, 2014, pp. 537– 548.
[4] Narayanan and V. Shmatikov, "De-anonymizing social networks," in Proceedings of the 2009 30th IEEE Symposium on Security and Privacy, ser. SP '09. Washington, DC, USA: IEEE Computer Society, 2009, pp. 173–187.
[5] Zhou, J. Pei, and W. Luk, "A brief survey on anonymization techniques for privacy preserving publishing of social network data," SIGKDD Explor.Newsl., vol. 10, no. 2, pp. 12–22, Dec. 2008.
[6] Mislove, B. Viswanath, K. P. Gummadi, and P. Druschel, "You are who you know: Inferring user profiles in online social networks," in Proceedings of the Third ACM International Conference on Web Search and Data Mining, ser. WSDM '10. New York, NY, USA: ACM, 2010, pp. 251–260.
[7] K. Jonghyuk Song, Jonghyuk Song, ―Inference attack on browsing history of twitter users using public click analytics and twitter metadata,‖ IEEE Transactions on Dependable and Secure Computing, 2014.

