International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 1, Issue 3, pp 1-4, December-2017

Inference Attack Prevention Model on Social Networking Application Data Sanitization Method on User's Profile

Veernala Sireesha

Assistant Professor, Department of Computer Science And Engineering, Malla Reddy College of Engineering for Women, Hyderabad -500100, Telangana, India

Abstract: Nowadays social media is becoming very popular and used for marketing as per users profile. But for this, social networking sites share user's data with other marketing companies and it is possible the third party companies can use user's private data. Significant factor in multimedia mobile systems is social network, where users can send their photos, videos and other media files. On the other hand, the information (e.g., user Bio, posts, etc.) on social media platforms shared usually reveals lots of users private information. That can be mined and mistreated for the malicious reasons. To tackle privacy concerns, privacy preserving mechanisms adopted by many social network service providers, e.g. hiding users profiles, anonym zing user identity, etc. As an attributes result from user profiles are usually set such that it could be accessed to prevent personal information outflow only by friends. To understand the hidden attributes to the numerous efficiency of current privacy protecting mechanisms different attacks have been proposed. Almost solutions are based on the social networking links along with users or their behaviors. The proposed work is an inference attack prevention model on social networking application to solve prevention problem. To prevent inference attack we proposed data sanitization method on user's profile.

Keywords: Inference attack, Data mining, social network privacy, Data Sanitization, Collective Inference.

1. INTRODUCTION

Social networking websites are virtual communities that foster interaction and encourages among associates of a group by permitting them to connect with other users, post personal data and link their personal profiles to others profiles. In most cases, membership is attained by registering as a user of that website in web community. Regularly interacting and visiting with people who use that website makes ones network solider. Though many social networking websites are release to anyone, who belong to a specific real world occupation or some are open only to people in certain age group. Members of social networking websites are communicated by posting weblogs, video and music stream, messages and chatting. Social networking sites members frequently link smaller communities within the interwork. Members of the social networking websites allow endorsing themselves and their comforts by posting individual profiles that contain enough information for others to determine if they are involved in associating with that person. An opponent of social networking claim that can be used to outbreak privacy and it contributes to grasping behavior. Many people are free with the information they post concerning themselves, those websites are frequently used to investigate s social habits and persons character. Social networks permit users to make public the details about themselves and to get connected with their friends. Some of the information is to be private when it revealed inside these networks. On released data to predict private information it is possible to use machine learning algorithms. In this paper, explore that how to launch inference attacks to predict private information by using released social networking data [1].

Data mining technique is an Inference attack that is used analyzing data by performing in order to gain knowledge illegitimately about database or subject. In social networking inference attack can be made on this type of data i.e. Profiles, Friends Connection Information, Messages, Groups Database (SQL injection) [2].

Database inference attack is one type of sql injection attack. The example below shows a derivate of inference attack (Error-based SQL injection). When the stacked condition is executed, it verifies if the current user is the system administrator by the database engine. If the condition is true, the statement forces the database to throw an error as division by zero by executing. Otherwise, a valid instruction is executed. If he sees a database error the attacker will be able to conclude the database is run by system administrator [3].

ISSN (online): 2581-3048
Volume 1, Issue 3, pp 1-4, December-2017

Nowadays data can be collected by different means as we have Internet of Things (IoT) in action. As we are in the era of smart cities vast amount of user profiles data is getting collected by the different sensors that are getting used in smart city concept, which is build with the help of IoT as a base. These data if not properly secured or used can become threat to the users privacy [4].

2. LITERATURE SURVEY

Sanitization Technique

Different problems related to the private information leakage in online social networks have been addressed. By using users details alone cannot give better predictability. So, for giving an better perdition friendship links can also be addressed. On using a simple local classification method does not get better result of collective inference. But result of combination from collective inference implication along with the individual results can reduce classifier accuracy in a greater amount by removing details and friendship links. Then sanitization technique can be used in many situations to remove sensitive information. To find sensitive attributes collective inference can be used. The effectiveness of inference attack on private information can be reduced by using the proposed sanitization methods [5].

Collective Method

Xin gaun and Yingshu, addressed two issues: (a) how exactly inference attack launch by third party users to predict private information of friends, and (b) There are efficient way to defend in opposition to such an attack to attain a preferred tradeoff privacy utility. For the first issue, show that utilizing collectively both link information and attribute can significantly increase prediction accuracy for private information. For the second issue, explore the dependence relationships for privacy attributes and public attributes. Based on these results, implement a collective method that doesn't take disadvantages of various data manipulating methods to gurantee user data sanitizing does not incur a worst impact on data utility. Using collective method, it is possible that effectively sanitize social media data prior to release. A solution is given for the two addressed issues are proven to be effective towards three real social dataset [6].

Link-Based Classification

Getoor compared a variety of link-based divisions including the spread of loopy belief, iterative division and field relaxation label. Facebook, BibSonomy and Flickr, They are not trying to clean or hide the name of graph data. However, their focus is on certain types of data, namely group membership taken and announced, that can be used as a way to improve the accuracy of the relationship. Their descriptive method of classification (as opposed to link-based or data- based) is a natural part of our data classification, as we will treat group membership details as other details, as we make a favorite movie or book. Jones and Soltren crawl Face book data and analyze usage patterns among Face book users will use information details and profile posts. Their paper focuses on mistakes especially within the face book platform. They do not analyze face- to-face user information and do not discuss attempts to read anonymous face book details [7].

3. METHODOLOGY

Decision Tree Algorithm

The accuracy of the trees is greatly affected by the decision to split the strategies. The cutting methods are not the same as the trees that go back and forth. Decision tree uses many algorithms to divide a node into more than one sub node. The homogeneity of successive sub-nodes increases with the formation of sub-nodes. We can say that the cleanliness of the area increases with respect to the changing direction. The olive tree algorithm separates the nodes from all available variables and selects the divisions that lead to many similar sub-nodes. Decision tree algorithm is a separator that is presented as a recurring division of the sample space. The solution tree algorithm consists of nodes that can form a tree with roots, meaning a target tree with a node called root with no inbound edges. The area around the outgoing edge is called the test or internal node. All other nodes are called terminals (also identified as leaves or nodes). In the decision tree, each internal node divides the model space into more than one space according to a random function of the input values.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 1, Issue 3, pp 1-4, December-2017

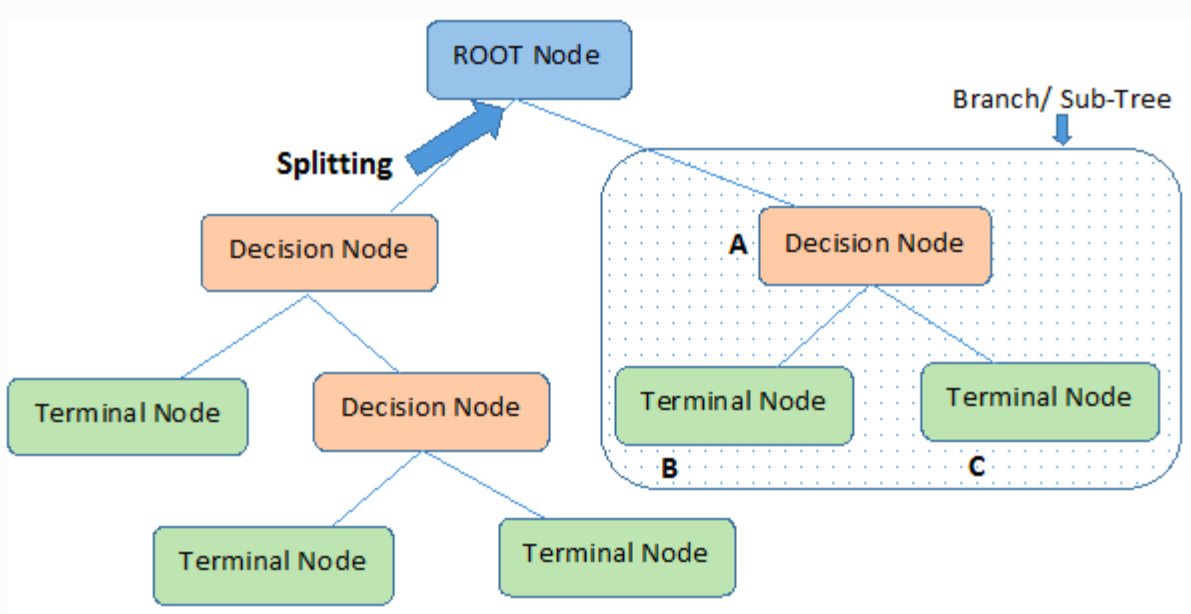


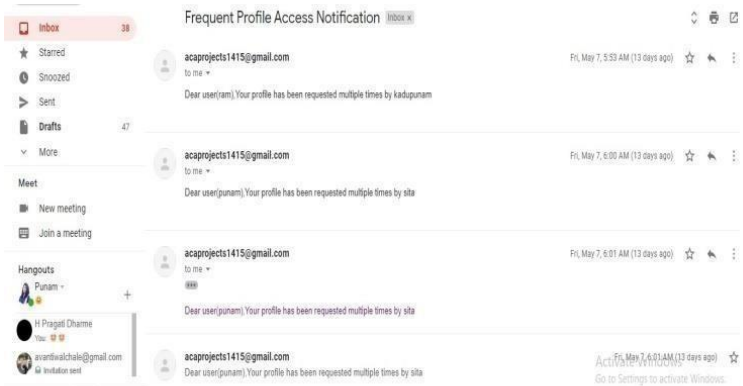
Fig 1: Decision Tree

AES Algorithm

AES is a pre-encrypted standard, used to encrypt data to keep it confidential. Protecting electronic data used is a cryptographic algorithm approved by FIPS. It is a blocking method that can encrypt and encrypt data. The encryption component converts data into a cipher text form while the transcription component converts the cipher text into text data form. A standard pre-encrypted algorithm used 128/192/256 bits to encrypt and encrypt data into 128-bit blocks. AES uses both software and hardware to protect digital data from multiple forms of data, voice, video, etc. from attack or audio.

Caesar Encryption Algorithm

In Encryption Details, Caesar cipher, also known as Caesars cipher, is the simplest and most well- known method of encryption. Cesari Encryption algorithm is a single type of text substitution for each punctuation mark that is replaced by a letter with a fixed number of letters under the letters of the alphabet. First we need to write down all the letters of the alphabet. It will now determine the encryption number. For example, it could be -1, -2, -3 etc or 1,2,3 etc. We will use +2 in this example. Now write all the benefits of the alphabet under the first one but move it 2 times to the right and move the remaining letters from end to end.

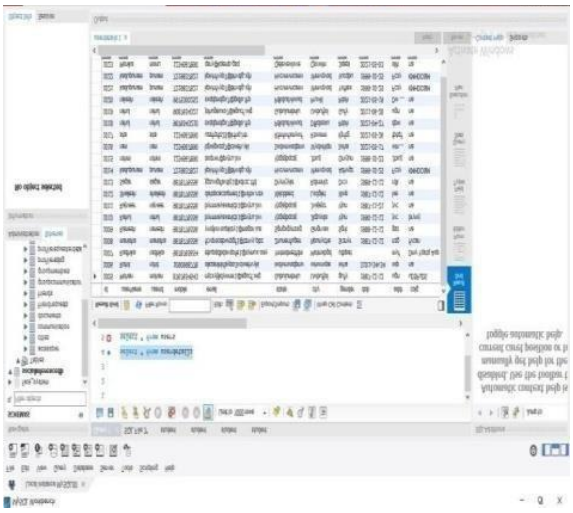


Output 1: Profile Access Notification

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 1, Issue 3, pp 1-4, December-2017



Output 2: Registered User's Database in Encrypted Format

4. CONCLUSION

We addressed two issues: (a) how exactly third party users launch an inference attack to predict private information of users, and (b) there are effective strategies to protect against such an attack to achieve a desired privacy utility tradeoff. For the first issue, show that collectively utilizing both link information and attribute can significantly increase prediction accuracy for sensitive information. For the second issue, explore the dependence, we propose a . Decision Tree Algorithm - This algorithm will be used to predict important data which we want to prevent. And Caesar Algorithm - This algorithm will be used for data encryption in database.

REFERENCES

[1] Z. Yin, M. Gupta, T. Weninger, and J. Han, "A unified framework for link recommendation using random walks," in Proc. IEEE Int. Conf. Adv. Social Netw. Anal. Mining. Odense, Denmark, Aug. 2010, pp. 152_159.

[2] Narayanan and V. Shmatikov, "De-anonym zing 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.

[3] J. He, W. W. Chu, and Z. V. Liu, "Inferring privacy information from social networks," in Proceedings of the 4th IEEE International Conference on Intelligence and Security Informatics, ser. ISI'06. Berlin, Heidelberg: Springer-Verlag, 2006, pp. 154– 165.

[4] AW. Burange and H. D. Misalkar, "Review of Internet of Things in development of smart cities with data management & privacy," 2015 International Conference on Advances in Computer Engineering and Applications, 2015, pp. 189-195.

[5] E. Zheleva and L. Getoor, "Preserving the Privacy of Sensitive Relationships in Graph Data," Proc. First ACM SIGKDD Int'l Conf. Privacy, Security, and Trust in KDD, pp. 153-171, 2008.

[6] Z. Jorgensen, T. Yu, and G. Cormode, "Publishing attributed social graphs with formal privacy guarantees," in Proceedings of the 2016 International Conference on Management of Data, ser. SIGMOD '16,2016, pp. 107–122.

[7] Barnaghi, P., Wang, W., Henson, C., and Taylor, K., "Semantics for the Internet of Things: Early Progress and Back to the Future," International Journal on Semantic Web and Information Systems, vol. 8, No. 1, 2012.

****

