PRIVACY-PRESERVING SOCIAL MEDIA DATA PUBLISHING FOR PERSONALISED RANKING BASED RECOMMENDATIONS

Mrs.Vijayalaksmi C[1] ,Pavithra R[2], Sharmila R[3,] Prathana Elizabeth Jyoti J[4]

[1]Assistant Professor,Dept. of CSE, Panimalar Engineering College, Chennai , India
[2,3,4] IV Yr Dept. of CSE, Panimalar Engineering College, Chennai , India

Abstract

In this paper, we proposed PrivRank, a customizable and continuous privacy-preserving social media data publishing framework protecting users against inference attacks while enabling personalized ranking-based recommendations. Its key thought is to ceaselessly jumble client action information with the end goal that the security spillage of client determined private information is limited under a given information bending spending plan, which limits the positioning misfortune brought about from the information jumbling process keeping in mind the end goal to save the utility of the information for empowering proposals.

Keywords— *PrivRank, DataMining,Big data, C#,*

I. INTRODUCTION

Developing effective recommendation engines is critical in the era of Big Data in order to provide pertinent information to the users. To deliver high-quality and personalized recommendations, online services such as e-commerce applications typically rely on a large collection of user data, particularly user activity data on social media, such as tagging/rating records, comments, check-ins, or other types of user activity data. In practice, many users are willing to release the data (or data streams) about their online activities on social media to a service provider in ex-change for getting high-quality personalized recommendations. In this paper, we refer to such user activity data as public data. However, they often consider part of the data from their social media profile as private, such as gender, income level, political view, or social contacts. In the following, we refer to those data as private

data. Although users may refuse to release private data, the inherent correlation between public and private data often causes serious privacy leakage. For example, one's political affiliation can be inferred from her rating of TV shows one's gender can be inferred from her activities on location-based social networks . These studies show that private data often suffers from inference attacks , where an adversary analyzes a user's public data to illegitimately gain knowledge about her private data. It is thus crucial to protect user private data when releasing public data to recommendation engines. To tackle this problem, privacy-preserving data publishing has been widely studied . Its basic idea is to provide i.e., whether the recommendation engines can accurately predict the individual's preference based on the obfuscated data. There is an intrinsic trade-off between privacy and personalization. On one hand, more distortion of public data leads to better privacy protection, as it makes it harder for adversaries to infer private data. On the other hand, it also incurs a higher loss in utility, as highly distorted public data prevents recommendation engines from accurately predicting users' real preferences.

II. LITERATURE SURVEY

A. *How to hide the elephant or the donkey in the room :Practical privacy against inference for large data*

This paper was submitted by S. Salamatian, A. Zhang, F. du Pin Calmon, S. Bhamidipati,N. Fawaz, B. Kveton, P. Oliveira, and N. Taft.The main objective is data is distorted before it is released, according to a privacy-preserving probabilistic mapping. This mapping is obtained by solving a convex optimization problem, which minimizes information leakage under a distortion constraint.

B. Privcheck: Privacy-preserving check-in data publishing for personalized location based services

This paper was submitted by D. Yang, D. Zhang, Q. Bingqing, in the year 2016. In this paper, we propose PrivCheck, a customizable and continuous privacy-preserving check-in data publishing framework providing users with continuous privacy protection against inference attacks. The key idea of PrivCheck is to obfuscate user check-in data such that the privacy leakage of user-specified private data is minimized under a given data distortion budget, which ensures the utility of the obfuscated data to empower personalized LBSs.

C. Protecting individual information against inference attacks in data publishing.

This paper was submitted by C. Li, H. Shirani-Mehr, and X. Yang in 2007.In many data-publishing applications, the data owner needs to protect sensitive information pertaining to individuals. Meanwhile, certain information is required to be published. The sensitive information could be considered as leaked, if an adversary can infer the real value of a sensitive entry with a high confidence. In this paper we study how to protect sensitive data when an adversary can do inference attacks using association rules derived from the data.

D. Privacy-preserving data publishing: A survey of recent developments

This paper was submitted by B. Fung, K. Wang, R. Chen, and P. S. Yu in 2010. Privacy Preserving Data Publishing (PPDP) is a way to allow one to share anonymous data to ensure protection against identity disclosure of an individual. Data anonymization is a technique for PPDP, which makes sure the published data, is practically useful for processing (mining) while preserving individuals sensitive information.

III. METHODOLOGY

The proposed system uses PrivRank, a customizable and continuous privacy-preserving social media data publishing framework protecting users against inference attacks while enabling personalized ranking-based recommendations. Its key idea is to continuously obfuscate user activity data such that the privacy leakage of user-specified private data is minimized under a given data distortion budget, which bounds the ranking loss incurred from the data obfuscation process in order to preserve the utility of the data for enabling recommendations.

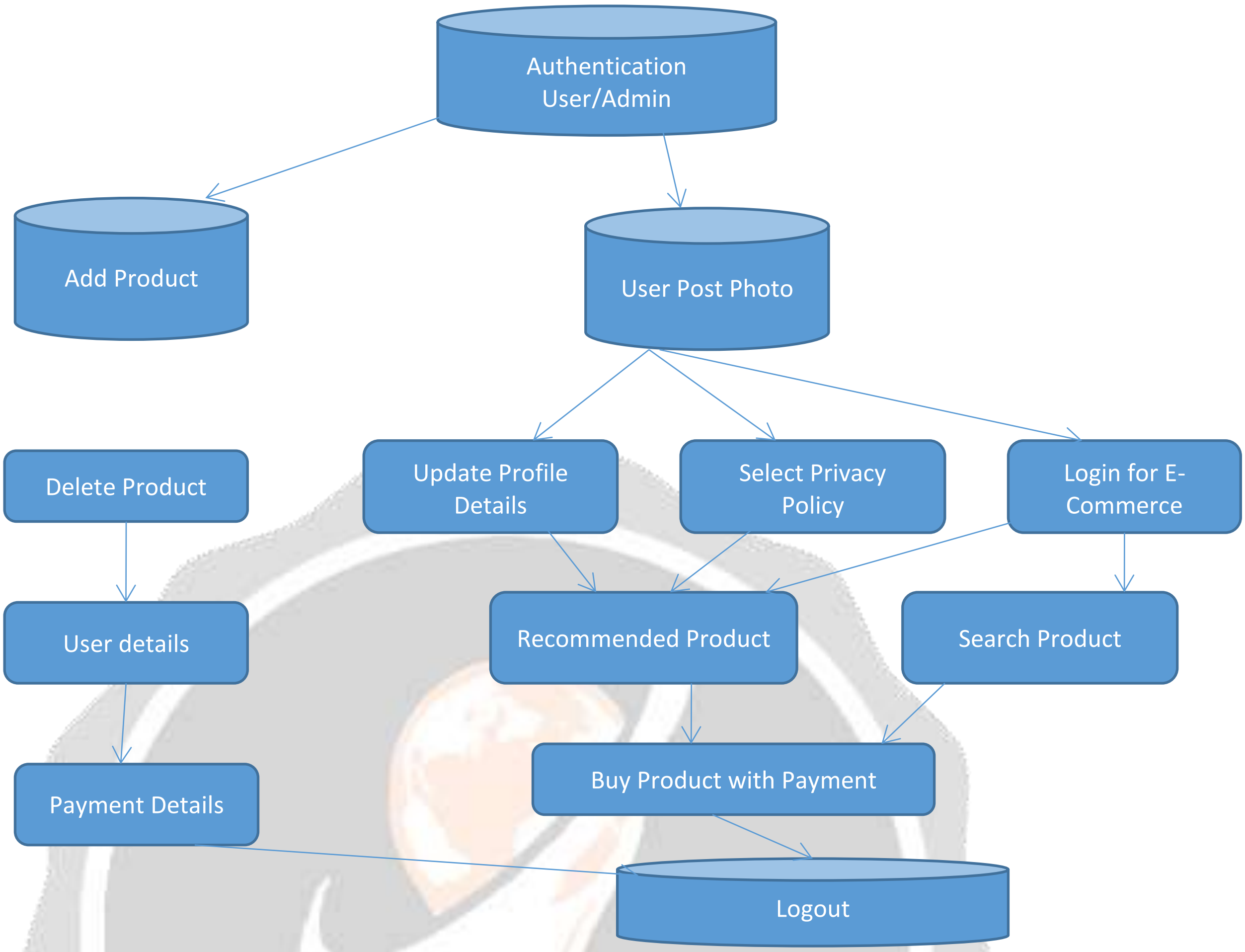

FIG 1 Overview of the system

Working:

a)User login-Registration

If you are a new user going to login into the application then you have to register first by providing necessary details. After successful completion of sign up process, the user has to login into the application by providing username and exact password. The necessary details given by the user will be saved for future purposes. The user needs to enter exact username and password. If login success means it will take up to upload page else it will remain in the login page itself. If login is not successful then an error will be displayed stating incorrect username and password.

b)Post status

In post status section, the user can post to access their social media and by uploading image to the Status with quote and hash tag .The posted status will be displayed in the current user's status. It can be modified if necessary and it will be stored in database for future reference. In update profile, the user has to update profile details like date of birth, gender, and hobbies. Also the user needs to select their interested topics such as gaming, sports, music, book and so on. The topics which we have selected here will be suggested in the recommended product.

C)Selecting privacy policy

Before buying the product the user needs to select one of the privacy policy. The three options for selecting privacy policy are Access all, Partially access and Don't access. If the user selects Access all then all the details that he gives will be accessed. If the user selects Partially access then only interested topics will be accessed. If the user selects Don't Access then a message will be displayed asking the user to select the privacy policy.

D)Recommended product

According to the users interest a list of items will be displayed in the recommended product. If the item which the user wants to buy does not appear in the recommended product then he can search for the product using keyword. User searching for a product with a key word and if the product is available with that key word user can select the product and buy it.

E)Buy product with payment

Once the user has selected the product, he can buy the product by giving payment details such as card number, card holder name, card type, expiry date, cvv and so on. Using these details the user can buy the product. The interested topics selected by the user previously will be shown in the recommended products. It checks for the privacy policy. Once the user has selected the privacy policy the products will show otherwise it asks for selecting one privacy policy.

F)Admin login

The Admin needs to enter exact username and password. If login success means it will take up to upload page else it will remain in the login page itself. Here the admin can add a product, delete a product and maintain user details. After the successful login the admin can add product where he can give description about the product, title of the product ,tag for the product ,image for the product and also he can give valid price for all the products added.All the details will be stored in the database. The admin maintains user details and user payment details for all process. It will be stored in the database as table view. The admin can make modifications if necessary.

IV.RESULT

To provide customized protection, the optimal data obfuscation is learned such that the privacy leakage of user-specified private data is minimized; to provide continuous privacy protection, we consider both the historical and online activity data publishing; to ensure the data utility for enabling ranking-based recommendation, we bound the ranking loss incurred from the data obfuscation process using the Kendall-_ rank distance.

V. CONCLUSION

We showed through extensive experiments that PrivRank can provide an efficient and effective protection of private data, while still preserving the utility of the published data for different ranking-based recommendation use cases.

VI. REFERENCES

[1] .S. Salamatian, A. Zhang, F. du Pin Calmon, S. Bhamidipati, N. Fawaz, B. Kveton, P. Oliveira, and N. Taft, “How to hide the elephant-or the donkey-in the room: Practical privacy against statistical inference for large data,” in Proc. of GlobalSIP. IEEE, 2013.

[2] D. Yang, D. Zhang, Q. Bingqing, and P. Cudre-Mauroux, “Privcheck: Privacy-preserving check-in data publishing for personalized location based services,” in Proc. of UbiComp’16. ACM, 2016.

[3] C. Li, H. Shirani-Mehr, and X. Yang, “Protecting individual information against inference attacks in data publishing,” in Advances in Databases: Concepts, Systems and Applications. Springer, 2007, pp. 422–433.

[4] B. Fung, K. Wang, R. Chen, and P. S. Yu, “Privacy-preserving data publishing: A survey of recent developments,” ACM Computer Survey, vol. 42, no. 4, p. 14, 2010.

[5] I. A. Junglas, N. A. Johnson, and C. Spitzm ¨ uller, “Personality traits and concern for privacy: an empirical study in the context of location-based services,” European Journal of Information Systems, vol. 17, no. 4, pp. 387–402, 2008.

[6] P. Cremonesi, Y. Koren, and R. Turrin, “Performance of recommender algorithms on top-n recommendation tasks,” in Proc. of RecSys’10. ACM, 2010, pp. 39–46.

[7] N. Li, R. Jin, and Z.-H. Zhou, “Top rank optimization in linear time,” in Advances in neural information processing systems, 2014, pp. 1502–1510.

[8] M. G. Kendall, “Rank correlation methods.” 1948.