WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

PRIVACY-PRESERVING SOCIAL MEDIA DATA PUBLISHING FOR PERSONALIZED RANKING-BASED RECOMMENDATION

Ms.M.Shamilie (ME) [1] , M.Rajeshwari [2], P.Nisha[3], M.Srivithya[4], A.Nandhini[5]

Assistant professor) [1] , Department of CSE, Final year CSE Students[2345]
Shreenivasa Engineering College ,Affiliated To Anna University ,Chennai
e-mail: rajim2820@gmail.com

ABSTRACT

Personalized recommendation is resolve to help users find suitable information. It generally confide on a mainly collection of user data, in appropriate users online activity (e.g., tagging/rating/checking-in) on social media, to mine user desire. In this paper, we consider PrivRank, a custom and continuous privacy-preserving social media data publishing framework covering users negative conclude attacks while entitle personalized ranking-based recommendations. An experimental evaluation on both artificial and real-world datasets display that our framework can skillfully provide effective and endless protection of user-specified private data, while still preserving the service of the complicate data for personalized ranking-based recommendation.

Keywords*: Privacy-preserving data publishing, personalized confidential protection, Personalization, Ranking-based recommendation and Social media.*

1. INTRODUCTION

Developing adequate recommendation engines is demanding in the era of Big Data in order to provide related information to the users. To deliver high-quality and personalized recommendations, online services such as browsing applications typically rely on a broad collection of user data, particularly user activity data on social media, such as marking/valuation enroll, comments, sign-in, or other types of user activity data. In process, many users are willing to release the data (or data streams) about their online activities on social media to a service provider in transfer for getting high-quality personalized recommendations.

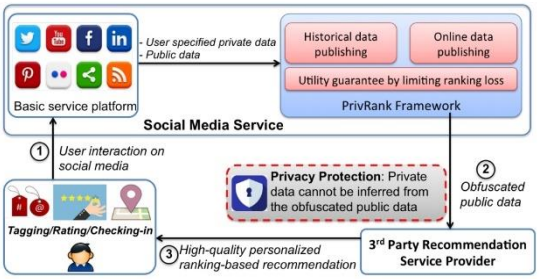
2. RELATED WORK

To defend user privacy when publishing user data, the latest practice mostly relies on action or user concurrency, e.g., on the use and magazine of the published data [4]. However, this approach cannot guarantee that the users sensitive information is actually protected from a cruel attacker. Therefore to provide effective privacy production when releasing user data, privacy preserving data publishing has been widely studied. Its key idea is to obfuscate user data such that published data remains useful for some application scenarios while the individual's privacy is preserved. comforting to the attacks designed, extant work can be confidential into two categories**.**

3. PRELIMINARIES

System Workflow:

Figure 1 explain the end-to-end workflow of our system. PrivRank is achieve as a increased module to existing social media platforms, in order to let user enjoy high-quality personalized recommendations from third party services under a personalized privacy guarantee.



User Preference Modeling from Social Media Data :

User's activities on social media densely imply their preferences. Particular social media services often provide users with a unique feature (or a certain type of items) for communication, such as photos for Flickr, videos for YouTube, music for Last.fm, and POIs for Foursquare. By collaborate with these items on social media (e.g., tagging a photo, rating a video or checking-in at a POI), users absolutely or essentially explicit their preferences on those items. In this work, we examine such user activity as mutual data. Correctly, let U and I stand for the sets of users and items, respectively.

Ranking-Based Recommendation:

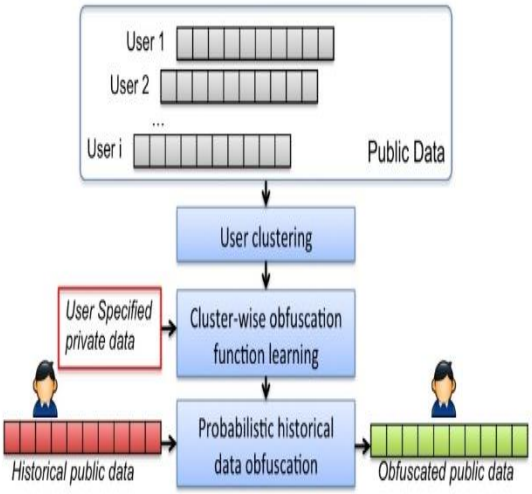
Based on the specified public data vectors, ranking based recommendation outputs a rated list of items for a user, where the top items are most likely to be tempting to her. The related algorithms mainly advantage the existing ranking of items in the learning process to anticipate the missing rank of the items for recommendation [16]. Therefore, ranking-based recommendation algorithms are susceptible to the ranking loss caught from the data confusion process, rather than other types of loss consistent by the Euclidean or Squared L 2 distance, for example. Moreover, those traditional data bias measures are not analogous to ranking loss [16]. Figure 2 shows an example where the same data bias budget consistent by Euclidean distance may signify different ranking losses. Therefore, considering ranking loss gain from data obfuscation is critical for ranking based recommendation.

4. THREAT MODEL

In this study, we consider the inference attack as the targeted threat model. As described above, we consider that each user has two types of data: i) public data (e.g., her activity data) that she is willing to release for getting personalized recommendations, and ii) private data (e.g., gender) that she wants to control confidential. We denote public data as $X \in X$, and private data as $Y \in Y$, where X and Y are the sets of values that X and Y can take, respectively. Since Y is often linked to X by their joint probability p(X, Y), an rival who observes X is able to gain some knowledge about Y . To reduce such privacy leakage, the basic idea is to release a distorted $X^{\wedge} \in X^{\wedge}$ instead of X such that it is hard to infer Y from $X^{\wedge}$.

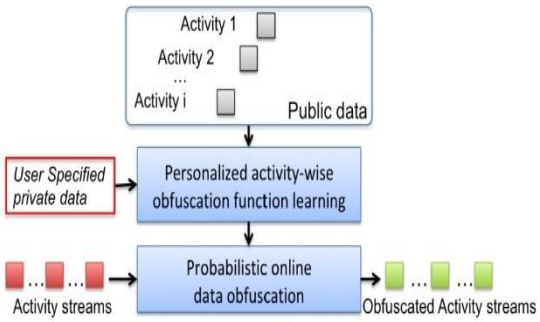
5. HISTORICAL DATA PUBLISHING

To publish actual public data in a privacy-preserving way, the key idea is to probabilistically complicate a user's historical public data vector to that of another user, which are similar but have less privacy leakage. In this context, data confusion operates on one's whole public data vector, rather than obfuscating her particular activity records one by one (over the user's activity stream). Compared to the effective scheme, we show that such a actual data confusion scheme can achieve the same level of privacy protection with a lower data bias budget.



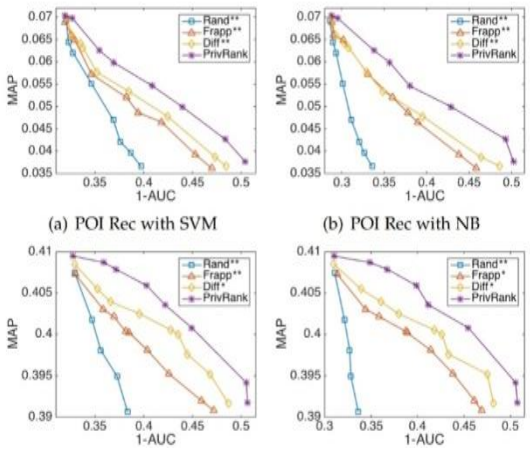
6. ONLINE DATA STREAM PUBLISHING

After a user joined to third-party services, the service providers have approach to the user's destiny activity streams. Therefore, we protect her confidential data by obfuscating her activity stream on-the-fly. Different from historical data publishing, the effective nature of user activity enforces the following constraint on online data obfuscation: Due to time and space ability requirements of real-time data publishing.



7. EXPERIMENTAL EVALUATION

We analytically evaluate the effectiveness and ability of our framework. Specifically, based on both synthetic and real-world datasets, we first investigate the trade-off between privacy protection and personalization performance for ranking-based recommendation. Second, we study the endless privacy protection performance by analyzing the privacy leakage over time.



8. CONCLUSIONS AND FUTURE WORK

This paper imported PrivRank, a justifiable and endless privacy-preserving social media data publishing framework. It frequently protects user-specified data against assumption attacks by releasing obfuscated user activity data, while still assure the utility of the released data to power personalized ranking-based recommendations.

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 obfuscation incurred from the data process using the Kendall-τ rank distance 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.

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.