Social Media Data Publishing Using Ranking Based Recommendation

Nandhini A[1], Jeevitha.T[2], Saran.R[3], Harish.C[4], Vignesh.M[5]
[1, 2]Asst. Professor, Dept of CSE
[3, 4, 5]Dept of CSE
[1]Sri Ramakrishna Institute of Technology, Coimbatore, India
[2, 3, 4, 5]Info Institute of Engineering, Coimbatore, India

Abstract- *In this project, we proposed Privacy Rank, an adaptable and nonstop security safeguarding internet based life information distributing structure ensuring clients against surmising assaults while empowering customized positioning based proposals. A social behavioral profile accurately reflects a user's social media activity patterns. While an authentic owner conforms to its account's social behavioral profile involuntarily, it is hard and costly for impostors to feign. In this project, we enhance the first computational mechanism to resolve conflicts for multi-party privacy management in Social network that is able to adapt to different situations by modeling the concessions that users make to reach a solution to the conflicts. Contrasted with cutting edge approaches, Privacy Rank accomplishes both a superior security assurance and a higher utility in all the positioning based suggestion use cases we tried.*

I. INTRODUCTION

User interaction with different online social media data publishing services, we propose several new behavioral features that can effectively quantify user differences in online social activities. To validate the effectiveness of social behavioral profile in detecting account activity anomaly, we apply the social behavioral profile of each user to differentiate click streams of its respective user from all other users. We categorize user social behaviors on an online social media data publishing into two classes, extroversive behaviors and introversive behaviors.

In this project, we proposed Privacy Rank, an adaptable and nonstop security safeguarding internet based life information distributing structure ensuring clients against surmising assaults while empowering customized positioning based proposals. A social behavioral profile accurately reflects a user's social media activity patterns. While an authentic owner conforms to its account's social behavioral profile involuntarily, it is hard and costly for impostors to feign. In this project, we enhance the first computational mechanism to resolve conflicts for multi-party privacy management in Social network that is able to adapt to different situations by modeling the concessions that users make to reach a solution to the conflicts. Contrasted with cutting edge approaches, Privacy Rank accomplishes both a superior security assurance and a higher utility in all the positioning based suggestion use cases we tried.

II. LITERATURE SURVEY

BPR: Bayesian Personalized Ranking from Implicit Feedback

Item recommendation is the task of predicting a personalized ranking on a set of items (e.g. websites, movies, products). In this paper, we investigate the most common scenario with implicit feedback (e.g. clicks, purchases). There are many methods for item recommendation from implicit feedback like matrix factorization (MF) or adaptive knearest-neighbor (kNN). Even though these methods are designed for the item prediction task of personalized ranking, none of them is directly optimized for ranking. In this paper we present a generic optimization criterion BPR-Opt for personalized ranking that is the maximum posterior estimator derived from a Bayesian analysis of the problem. We also provide a generic learning algorithm for optimizing models with respect to BPR-Opt. The learning method is based on stochastic gradient descent with bootstrap sampling. We show how to apply our method to two state-of-the-art recommender models: matrix factorization and adaptive kNN. Our experiments indicate that for the task of personalized ranking our optimization method outperforms the standard learning techniques for MF and kNN. The results show the importance of optimizing models for the right criterion.

Protecting Individual Information Against Inference Attacks in Data Publishing

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. We formulate the inference attack model, and develop complexity results on computing a safe partial table. We classify the general problem into subcases based on the requirements of publishing information, and propose the corresponding algorithms for finding a safe partial table to publish. We have conducted an empirical study to evaluate these algorithms on real data.

A Sentiment-Enhanced Personalized Location Recommendation System

Although online recommendation systems such as recommendation of movies or music have been systematically studied in the past decade, location recommendation in Location Based Social Networks (LBSNs) is not well investigated yet. In LBSNs, users can check in and leave tips commenting on a venue. These two heterogeneous data sources both describe users' preference of venues. However, in current research work, only users' check-in behavior is considered in users' location preference model, users' tips on venues are seldom investigated yet. Moreover, while existing work mainly considers social influence in recommendation, we argue that considering venue similarity can further improve the recommendation performance. In this research, we ameliorate location recommendation by enhancing not only the user location preference model but also recommendation algorithm. First, we propose a hybrid user location preference model by combining the preference extracted from check-ins and text-based tips which are processed using sentiment analysis techniques. Second, we develop a location based social matrix factorization algorithm that takes both user social influence and venue similarity influence into account in location recommendation. Using two datasets extracted from the location based social networks Foursquare, experiment results demonstrate that the proposed hybrid preference model can better characterize user preference by maintaining the preference consistency, and the proposed algorithm outperforms the state-of-the-art methods.

Modeling User Activity Preference by Leveraging User Spatial Temporal Characteristics in LBSNs

With the recent surge of location based social networks (LBSNs), activity data of millions of users has become attainable. This data contains not only spatial and temporal stamps of user activity, but also its semantic information. LBSNs can help to understand mobile users' spatial temporal activity preference (STAP), which can enable a wide range of ubiquitous applications, such as personalized context-aware location recommendation and group-oriented advertisement. However, modeling such user-specific STAP needs to tackle high-dimensional data, i.e., user-location-time-activity quadruples, which is complicated and usually suffers from a data sparsity problem. In order to address this problem, we propose a STAP model. It first models the spatial and temporal activity preference separately, and then uses a principle way to combine them for preference inference. In order to characterize the impact of spatial features on user activity preference, we propose the notion of personal functional region and related parameters to model and infer user spatial activity preference. In order to model the user temporal activity preference with sparse user activity data in LBSNs, we propose to exploit the temporal activity similarity among different users and apply nonnegative tensor factorization to collaboratively infer temporal activity preference. Finally, we put forward a context aware fusion framework to combine the spatial and temporal activity preference models for preference inference. We evaluate our proposed approach on three real-world datasets collected from New York and Tokyo, and show that our STAP model consistently outperforms the baseline approaches in various settings.

PrivCheck: Privacy-Preserving Check-in Data Publishing for Personalized Location Based Services

With the widespread adoption of smartphones, we have observed an increasing popularity of Location-Based Services (LBSs) in the past decade. To improve user experience, LBSs often provide personalized recommendations to users by mining their activity (i.e., check-in) data from location-based social networks. However, releasing user check-in data makes users vulnerable to inference attacks, as private data (e.g., gender) can often be inferred from the users' check-in data. 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. Since users often give LBS providers access to both their historical check-in data and future check-in streams, we develop two data obfuscation methods for historical and online check-in publishing, respectively. An empirical evaluation on two real-world datasets shows that our framework can efficiently provide effective and continuous protection of user-specified private data, while still preserving the utility of the obfuscated data for personalized LBSs.

Privacy-Preserving Personalized Recommendation: An Instance-based Approach via Differential Privacy

Recommender systems become increasingly popular and widely applied nowadays. The release of users' private data is required to provide users accurate recommendations, yet this has been shown to put users at risk. Unfortunately, existing privacy preserving methods are either developed under trusted server settings with impractical private recommender systems or lack of strong privacy guarantees. In this paper, we develop the first lightweight and provably private solution for personalized recommendation, under untrusted server settings. In this novel setting, users' private data is obfuscated before leaving their private devices, giving users greater control on their data and service providers less responsibility on privacy protections. More importantly, our approach enables the existing recommender systems (with no changes needed) to directly use perturbed data, rendering our solution very desirable in practice. We develop our data perturbation approach on differential privacy, the state-of-the-art privacy model with lightweight computation and strong but provable privacy guarantees. In order to achieve useful and feasible perturbations, we first design a novel relaxed admissible mechanism enabling the injection of flexible instance-based noises. Using this novel mechanism, our data perturbation approach, incorporating the noise calibration and learning techniques, obtains perturbed user data with both theoretical privacy and utility guarantees. Our empirical evaluation on large-scale real-world datasets not only shows its high recommendation accuracy but also illustrates the negligible computational overhead on both personal computers and smartphones. As such, we are able to meet two contradictory goals, privacy preservation and recommendation accuracy. This practical technology helps to gain user adoption with strong privacy protection and benefit companies with high-quality personalized services on perturbed user data.

PriView: Media Consumption and Recommendation Meet Privacy Against Inference Attacks

We propose PriView, an interactive privacy preserving personalized video consumption system, that protects a user's privacy while delivering relevant content recommendations to the user. PriView provides the user with three functionalities: transparency on privacy risk, control of privacy risk, and personalized content recommendations. PriView bridges privacy theory and practice: it successfully implements an information theoretic framework to design a utility-aware privacy-preserving mapping that perturbs a user's video ratings to prevent inference of user private attributes, e.g. political views, age, gender, while maintaining the utility of the released perturbed ratings for recommendation. Our model uses convex optimization to learn a probability mapping from actual ratings to perturbed ratings that minimizes distortion subject to a privacy constraint. One practical challenge of the optimization is scalability, when the size of the underlying alphabet of the user data is very large, e.g. due to a large number of features representing the data. To reduce the optimization size, we introduce a quantization step that allows to control the number of optimization variables, and explore using low rank approximations of the rating matrix. Evaluations on the Politics and TV dataset show that these methods can achieve perfect privacy with little chan

ge in recommendation quality.

K-Anonymity: A Model For Protecting Privacy

Consider a data holder, such as a hospital or a bank, that has a privately held collection of person-specific, field structured data. Suppose the data holder wants to share a version of the data with researchers. How can a data holder release a version of its private data with scientific guarantees that the individuals who are the subjects of the data cannot be re-identified while the data remain practically useful? The solution provided in this paper includes a formal protection model named fc-anonymity and a set of accompanying policies for deployment. A release provides fc-anonymity protection if the information for each person contained in the release cannot be distinguished from at least *k*- individuals whose information also appears in the release.

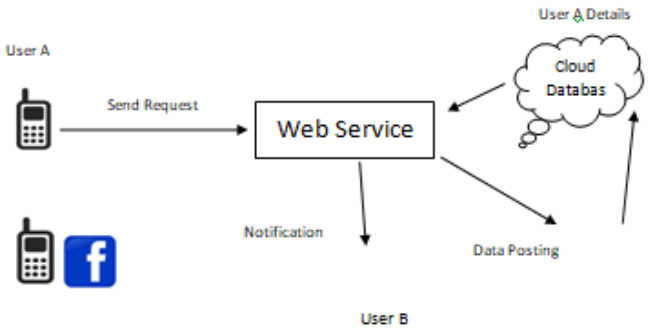
Checking in without Worries: Location Privacy in Location Based Social Networks

In current location based social networks (LBSNs), users expose their location when they check in at a venue or search a place. The release of location privacy could lead to a severe breach of other privacy, such as identity or health condition. In this paper, we propose a framework to safeguard users' location information as well as the check-in records. Considering the special demands in LBSNs, we design a novel index structure to provide a fast search for users when they check in at the same venue frequently. At the same time, our framework outsources the heavy cryptographic computations to the server to reduce the computational overhead for mobile clients. Due to the dynamic feature of LBSNs, our framework uses a lightweight approach to handle a user's revoked friends and new friends. We prove the security of our framework in the random oracle model and demonstrate its efficiency on a Motorola Droid phone.

Utility-Privacy Tradeoff in Databases: An Information-theoretic Approach

Ensuring the usefulness of electronic data sources while providing necessary privacy guarantees is an important unsolved problem. This problem drives the need for an analytical framework that can quantify the privacy of personally identifiable information while still providing a quantifiable benefit (utility) to multiple legitimate information consumers. This paper presents an information-theoretic framework that promises an analytical model guaranteeing tight bounds of how much utility is possible for a given level of privacy and vice-versa. Specific contributions include: i) stochastic data models for both categorical and numerical data; ii) utility-privacy tradeoff regions and the encoding (sanization) schemes achieving them for both classes and their practical relevance; and iii) modeling of prior knowledge at the user and/or data source and optimal encoding schemes for both cases.

III. BLOCK DIAGRAM



IV. CONCLUSION

In Our Method proposed Advanced PrivRank mechanism with intrusion detection system, a customizable and continuous privacy-preserving social media data publishing framework with securely. It continuously protects user-specified data against inference attacks by releasing obfuscated user activity data, while still ensuring 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 with opinion mining; to ensure the data utility for enabling ranking-based recommendation. 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.

In the future, we plan to extend our framework by considering the data types with continuous values rather than discredited values, and explore further data utility beyond personalized recommendation.

REFERENCES

[1] Elias M.Award, ”System Analysis and Design”, Galgotia Publications, Second Edition.
[2] Robin Dewson, “Pro SQL Server 2005**”,** Apress Publisher.
[3] Paul Kimmel, “Asp.Net-Unleased”, Sams TechMedia, Third Edition.
[4] Roger S. Pressman, ”Software Engineering”, Fourth Edition, 2005.
[5] Steven Holzner, “Asp.Net-Black Book”, DreamTech Press Limited, Second Edition.
[6] http://www.asp.net
[7] http://www.osborne.com
[8] http://www.microsoft.com