POSTER: Recommending Location Privacy Preferences in Ubiquitous Computing

Yuchen Zhao
University of St Andrews
yz39@st-andrews.ac.uk

Juan Ye
University of St Andrews
jy31@st-andrews.ac.uk

Tristan Henderson
University of St Andrews
tnhh@st-andrews.ac.uk

ABSTRACT

Location-Based Services have become increasingly popular due to the prevalence of smart devices. The protection of users' location privacy in such systems is a vital issue. Conventional privacy protection methods such as manually predefining privacy rules or asking users to make decisions every time they enter a new location may not be usable, and so researchers have explored the use of machine learning to predict preferences. Model-based machine-learning classifiers which are used for prediction may be too computationally complex to be used in real-world applications. We propose a location-privacy recommender that can provide users with recommendations of appropriate location privacy settings through user-user collaborative filtering. We test our scheme on real world dataset and the experiment results show that the performance of our scheme is close to the best performance of model-based classifiers and it outperforms model-based classifiers when there are no sufficient training data.

Categories and Subject Descriptors

K.4.1 [**COMPUTERS AND SOCIETY**]: Public Policy Issues—*Privacy*

General Terms

Security

Keywords

Location-Based Services, privacy protection, recommender system

1. INTRODUCTION

The popularity of mobile devices such as smartphones and tablets makes computing and services accessible anytime and anywhere. In this ubiquitous computing environment, users' location information has become a new feature which prospers increasing numbers of Location-Based Services (LBSs) applications, e.g., My Track[1] and Foursquare[2]. LBSs personalise users' service experiences and enable them to share their location data and personal tracks to others. However, they introduce location privacy issues at the same time. Location data are highly sensitive since overexposed location information would lead to disclosure of private facts such as occupations, health conditions and sexual orientations, etc. The dissemination of disclosed private information may lead to more serious threats such as blackmail and damage of reputation. Thus the protection of location privacy is a vital issue.

Users can control their location data exposure by manually predefining privacy rules or being requested for permissions every time. These methods, however, are inconvenient and cumbersome for users [4]. Meanwhile, untrained users find it difficult to configure location privacy rules by themselves.

In order to solve the usability issues, model-based classifiers have been introduced to predict users' privacy preferences [1]. Model-based classifiers can build users' personal privacy preference models from their privacy decision histories and use these models to predict their privacy preferences and automatically configure privacy rules. The expense of training and updating models, however, is time consuming. It also suffers from cold start problems, which is the poor performance of prediction when a new user using the system without sufficient personal data to train the model. We are therefore interested in building a light-weight location privacy recommender which need not build models for prediction and can overcome the lack of personal information during cold start periods. Xie et al. [5] successfully apply CF for privacy recommendations by testing it on crowdsourced Amazon Mechanical Turk data. By contrast, we test CF on real world data set and analyse its performance during the cold start period.

2. METHOD

To provide usable location privacy protection in a light-weight way and overcome the cold start problem at the same time, we propose a location privacy recommender using user-user collaborative filtering (CF). The user-user CF [3] is a technique of recommender systems that has been used in many areas including electronic commerce and news feeds. It can predict the target user's ratings for unknown items based on opinions of other users who have similar previous ratings with the target user. The intuition behind our scheme is that users have similar privacy preferences in some contexts may also make similar decisions in other contexts, which is using the social choices to help users configure their personal privacy settings. The benefit of using user-user CF to predict privacy preferences is that it need not train personal models and can get information from others when there are no sufficient personal data.

[1] http://www.google.com/mobile/mytracks

[2] http://foursquare.com

WiSec'14, Jul 23-25 2014, Oxford, United Kingdom
ACM 978-1-4503-2972-9/14/07.

We introduce it as a light-weight scheme to recommend privacy preferences and also use it to overcome the cold start problems.

In user-user CF, the basic elements are users, items and ratings. To deploy it in privacy preference prediction, we use context (the combination of time and location) as items and transform users' location privacy histories to ratings to different contexts. When users seldom share their location in a specific context, they have low ratings for this context and vice versa. By this means, we have their privacy preferences described by a vector of ratings for all contexts and then calculate and compare their privacy preference similarities. When target users need privacy recommendations, we find the neighbours who have the highest privacy preference similarities with them and combine their opinions as the recommendation.

3. RESULTS

We use the LocShare dataset [2], which contains real-world user location privacy preference data collected in St Andrews. We use users' identifications ($N = 40$), categories of location, times when they made privacy decisions and the decisions in our experiment.

We aim to answer two research questions: (1) can user-user CF perform as well as model-based classifiers do? (2) can user-user CF overcome the cold start problem in predictions?

For the first research question, we use 10-fold cross-validation to test our scheme, model-based classifiers (Naive Bayes, J48 and Rotation Forest) and semantic crowdsourcing prediction. We consider two metrics: prediction accuracy (where the recommender predicts the same decision as the test data) and privacy leaks (where the recommender predicts sharing but the actual decision in the test data is to not share). Our results indicate that our scheme (73.00% accuracy, 11.82% leaks) outperforms semantic crowdsourcing prediction methods (55.68% accuracy, 21.00% leaks). Compared with model-based classifiers, the performance of our scheme is close to the best performance (Rotation Forest) (75.30% accuracy, 12.70% leaks), which is the most accurate, but also too expensive to be used in practice due to the cost of model training process.

For the second research question, to test the influence of the cold start stage, for each user, we use small fractions (from 1% to 10%) of personal data combined with entire sets of other users' data for training, and use the rest (from 99% to 90%) of personal data for testing. As shown in Figures 1 and 2, during the cold start stage, our scheme can provide higher prediction accuracy (except 4%) until using 6% of personal data for training. Our scheme causes less privacy leaks in the whole period than the best performance of model-based classifiers. It also outperforms the semantic crowdsourcing prediction both in terms of prediction accuracy and reducing privacy leaks.

4. CONCLUSION

Our experimental results suggest that location privacy recommenders by using user-user CF can perform as well as model-based classifiers do and can also cause fewer privacy leaks. Because of the elimination of model training process, our scheme is more light-weight than model-based classifiers to be used and can overcome the cold start problem, which is very common in real world applications. Compared with semantic crowdsourcing predictions, our scheme can make the prediction more personalised which leads to a better performance.

To better understand common users' acceptance of the recommenders, we plan to conduct user studies to investigate under what circumstances they trust the recommendations from social choices. We are also interested in if the form of recommendations (e.g., recommendation only, recommendation plus confidence or reasons)

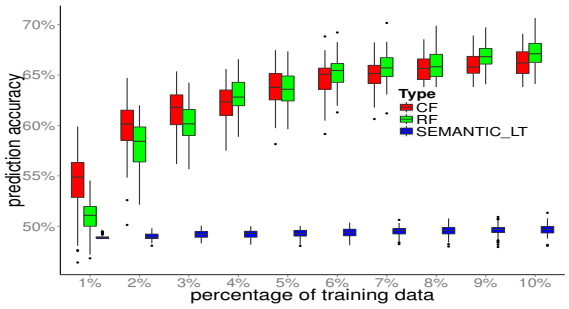

Figure 1: Prediction accuracies of CF, Rotation Forest (RF) and Location-Time Semantic (SEMANTIC_LT) during the cold start stage.

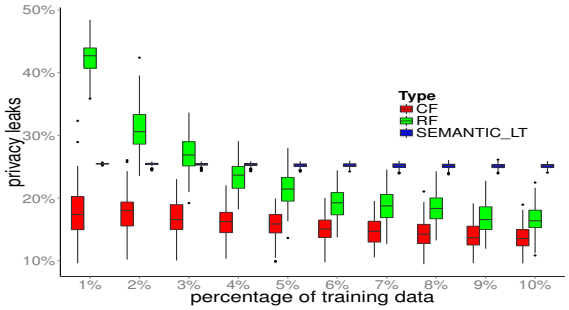

Figure 2: Privacy leaks of CF, Rotation Forest (RF) and Location-Time Semantic (SEMANTIC_LT) during the cold start stage.

has influences on users' decisions. The expected results will help us make the recommender more informative for users.

Users' privacy preferences are sensitive information, hence revealing their preferences to untrusted service providers may cause privacy implications. Our future work is to investigate how to enable users to use privacy recommenders in a privacy-aware way. Due to the elimination of personal models, a potential benefit of using user-user CF is when getting recommendations from crowdsourcing, users can do it without revealing their identities. To achieve that, mechanisms to protect the service from abusing or biasing by malicious users are necessary.

5. REFERENCES

[1] G. Bigwood, F. Ben Abdesslem, and T. Henderson. Predicting location-sharing privacy preferences in social network applications. In *Proc. of AwareCast*, 2012.

[2] I. Parris and F. Ben Abdesslem. CRAWDAD data set st_andrews/locshare (v. 2011-10-12). Downloaded from http://crawdad.org/st_andrews/locshare, Oct. 2011.

[3] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl. GroupLens: an open architecture for collaborative filtering of netnews. In *Proc. of CSCW*, pages 175–186, 1994.

[4] J. Y. Tsai, P. G. Kelley, L. F. Cranor, and N. Sadeh. Location-sharing technologies: Privacy risks and controls. In *In Proc. of TPRC*, 2009.

[5] J. Xie, B. P. Knijnenburg, and H. Jin. Location sharing privacy preference: analysis and personalized recommendation. In *Proc. of IUI*, pages 189–198, 2014.