Recommending Location Privacy Preferences in Ubiquitous Computing

University of St Andrews | FOUNDED 1413

Yuchen Zhao, Juan Ye and Tristan Henderson {yz39,jy31,tnhh}@st-andrews.ac.uk

Introduction

People share their locations in an increasing number of ways, such as checking in on Foursquare and Facebook. Peoples' location information is highly sensitive and inappropriate location exposure may cause privacy violations. Current mechanisms for preserving privacy suffer from usability issues:

- People find it difficult to configure location privacy rules appropriately [4].
- Model-based machine learning techniques [1] have been widely applied to help people make decisions, but they are computationally complex to use and suffers from the cold-start problem, that is, poor performance when there are insufficient training data.

We are therefore interested in building light-weight location privacy recommenders by using user-user collaborative filtering (CF) [3], which need not build models for prediction and can overcome the lack of personal information during the cold-start stage.

Method

To evaluate our system, we used the LocShare dataset [2], which comprises 3,878 location-sharing decisions made by 40 people in St Andrews. We divide the time of day into 5 slots and locations into 6 categories.

We denote the set of location attributes by L and the set of time attributes by T. We then use the cartesian product of T and L to represent the set of $item$s by:

$$I = T \times L$$

For each participant, we use his or her most frequent decision as the rating to the $item$ representing the time-location pair. We use the cosine similarity as the user similarity function.

We use the decision from the group whose group weight (i.e., the sum of similarities of users in the group) is higher than its counterpart as the prediction for the target user.

Results

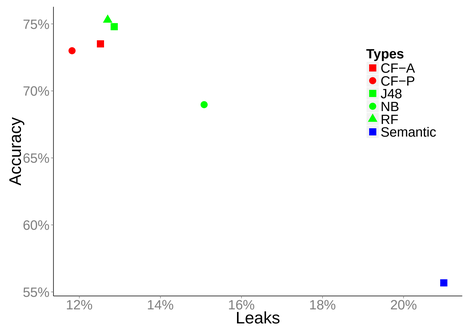
The following confusion matrix shows the possible prediction results and the actual decisions in the test set.

Actual Decision \ Prediction Preference	1	0
1	True Positive (TP)	False Negative (FN)
0	False Positive (FP)	True Negative (TN)

We define our metrics $Accuracy$ and $Leaks$ as:

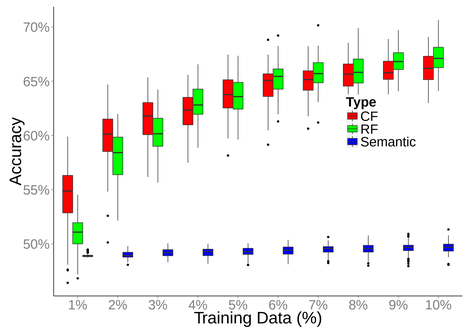
$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Leaks = \frac{FP}{TP + TN + FP + FN}$$



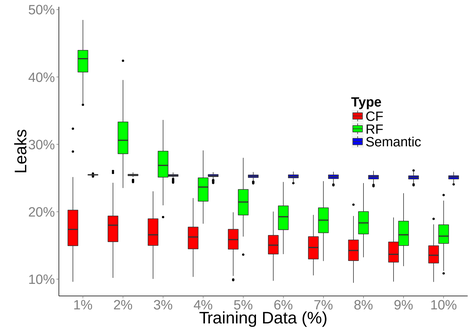
$Accuracy$ and $Leaks$ of CF (highest $Accuracy$ and lowest $Leaks$), model-based machine learning classifiers (J48, Naive Bayes, Rotation Forest) and semantic crowdsourcing prediction (using crowd preferences for the same location-time categories as the prediction [5]).

Our scheme outperforms semantic crowdsourcing prediction methods in terms of both $Accuracy$ and $Leaks$. The $Accuracy$ of using CF is close to the best performance of model-based classifiers and it causes fewer $Leaks$.



$Accuracy$ of CF, Rotation Forest (RF) and semantic crowdsourcing prediction during the cold-start stage.

During the cold start stage, our scheme can provide higher $Accuracy$ than RF (except 4%) until using 6% of personal data for training. The $Accuracy$ of using CF is higher than using semantic crowdsourcing prediction methods.



$Leaks$ of CF, Rotation Forest (RF) and semantic crowdsourcing prediction during the cold-start stage.

Our scheme causes fewer $Leaks$ than RF and semantic crowdsourcing prediction during the cold-start stage.

Future Research Questions

- Will people accept our system? Under what circumstances will they trust the recommendations from social choices rather than their own decisions?
- Will more informative feedback, such as displaying confidences or reasoning behind recommendations, help?
- Can we enable people to receive useful recommendations without revealing their real privacy preferences (e.g., obfuscating some sensitive ratings) to untrusted service providers?
- Can we prevent recommendation from being biased by malicious users or dishonest recommenders?

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. 2013.

[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] N. Sadeh, J. Hong, L. Cranor, I. Fette, P. Kelley, M. Prabaker, and J. Rao. Understanding and capturing people's privacy policies in a mobile social networking application. *Personal and Ubiquitous Computing*, 13(6):401–412, Oct. 2008.

[5] E. Toch. Crowdsourcing privacy preferences in context-aware applications. *Personal and Ubiquitous Computing*, 18(1):129–141, Dec. 2012.