ISSN 1913-8989 E-ISSN 1913-8997
Published by Canadian Center of Science and Education

Improved Recommender for Location Privacy Preferences

Anas A. Hadi[1] & Jonathan Cazalas[1]

[1] College of Computing and Information Technology, King Abdul-Aziz University, Jeddah, KSA

Correspondence: Anas A. Hadi, College of Computing and Information Technology, King Abdul-Aziz University, Jeddah, KSA. E-mail: anas1401@gmail.com

Received: October 23, 2015 Accepted: November 3, 2015 Online Published: November 6, 2015

doi:10.5539/cis.v8n4p64 URL: http://dx.doi.org/10.5539/cis.v8n4p64

Abstract

Location-based services are one of the fastest growing technologies. Millions of users are using these services and sharing their locations using their smart devices. The popularity of using such applications, while enabling others to access user's location, brings with it many privacy issues. The user has the ability to set his location privacy preferences manually. Many users face difficulties in order to set their preferences in the proper way. One solution is to use machine learning based methods to predict location privacy preferences automatically. These models suffer from degraded performance when there is no sufficient training data. Another solution is to make the decision for the intended user, depending on the collected opinions from similar users. *User-User Collaborative Filtering (CF)* is an example within this category. In this paper, we will introduce an improved machine learning based predictor. The results show significant improvements in the performance. The accuracy was improved from 75.30% up to 84.82%, while the privacy leak was reduced from 11.75% up to 7.65%. We also introduced an integrated model which combines both machine learning based methods and collaborative filtering based methods in order to get the advantages from both of them.

Keywords: location Based Services (LBS), mobile computing, preferences, privacy recommender

1. Introduction

With the huge evolution in the communication domain and the wide spread of smart devices, location based services has become very popular. Nowadays, millions of users are communicating through location-sharing applications. These applications include Facebook, Foursquare, Qype, Loopt, Snapchat, and so on.

Location based services provides the end user with many useful applications. Users have the ability to share their locations or path data with other users. But these applications, when associated with private user information, may raise serious privacy concerns (Wernke, Skvortsov, Dürr, & Rothermel, 2014). Private places, sensitive information, and empty homes, could be revealed, if location information is analyzed using time information. Analyzing location information and paths could end up with 60m approximation of home place (Krumm, 2007).

Location based services applications provides users with the ability of setting his privacy preferences manually. Many users face some difficulties in order to set their preferences in the proper way (Sadeh et al., 2009). On the other hand, the context has a major impact on selecting the appropriate preferences (Anthony, Henderson, & Kotz, 2007). Thus, it is useful to help the end user to select and refine his preferences.

Machine Learning based methods were proposed to overcome this issue. Machine learning methods learn from the previous preferences of the user and try to predict the future ones.

The training phase in these methods is essential and depends on the history of the user's preferences. The performance of machine learning methods is sensitive to the leak in training data amount. When the system has a new user, there won't be much data available to this user in order to train the system. This is called the "cold start period".

In order to deal with the problem of insufficient training data, Crowdsourcing based methods were proposed (Jin, Saldamli, Chow, & Knijnenburg, 2013; Toch, 2014). To make a decision, these methods depend on gathering the data from other users rather than the data history of the intended user.

Recommendation based methods are another variation of Crowdsourcing, where the recommender make the decision based on opinions from a group of similar users. *User-User collaborative filtering (CF)* is one good example of these methods. *User-User CF* depends on the concept whereby it is assumed that if there is a similarity between the users for the past preferences, then they may be a similarity between them for the future ones.

(Zhao, Ye, & Henderson, 2014) had successfully applied *User-User CF* as Location privacy preferences recommendation. The main advantage of using User-User CF is the ability to perform well, even under the condition of insufficient training data "cold start period". Their results were comparable to the performance of Machine Learning based models in general and better in the cold start period. Another advantage is that they provide a privacy-aware model which maintains the privacy of the users by obfuscating their preferences.

The contributions of this paper can be summarized as follows:

- ✓ First, we will apply *Random Forest Tree (RF)* as a Machine Learning method with some additional feature in order to enhance the prediction performance using real world dataset.
- ✓ Then we will discuss the positive and negative aspects of using both *User-User CF* and *Random Forest Tree (RF)*.
- ✓ Finally we will combine both *User-User CF* and *Random Forest Tree (RF)* in one integrated model.

The remainder of the paper is organized as follows: Section 2 will provide the related work and the necessary background. Section 3 will be devoted for the proposed work. While the results and discussions are covered in section 4. Finally, we will conclude the paper in the Conclusion section.

2. Realted Work

Millions of users with smart devices are attracted by Location Based Services (LBS) technology. But since the introduction of LBS, users' locations privacy emerged as an important issue. Usually the information required by LBS has the form of <User ID, Position, Time>. Thus, the main idea behind this privacy issue is to protect: *User identity, spatial information, or temporal information* (Wernke et al., 2014). Protecting the three of them - or at least two – is necessary. For example, if we hide the ID of the user, then the attacker may analyze both spatial information and temporal information to figure out user's ID.

2.1 Privacy Approaches

Many approaches have been suggested for Location privacy issue. We can summarize these approaches as:

2.1.1 Position Dummies

The main goal here is to protect the true position of the user by sending it among some other fake positions. These fake positions are called *dummies*. Usually, these approaches need a Trusted Third Party (TTP). *SybilQuery* (Shankar, Ganapathy, & Iftode, 2009) is an example of this approach.

2.1.2 Mixed Zones

The idea of this approach is to define an area or a zone where the true positions of the users are hidden inside this zone. This is done by not sending any position update while the user is still inside the same zone. Once the user has entered a zone, his ID will by mixed with the ID's of other users already in the zone. This approach was applied successfully to the road network in MobiMix (Palanisamy, & Li, 2011).

2.1.3 K-Anonymity

The concepts here is that the true position of the user can't be determined among other k-1 users. According to this, the probability of identifying the true user is $1/k$. Trusted Third Party (TTP) is needed to anonymize the user among other users. The user can define a K value and the area space (Mokbel, Chow, & Aref, 2006). K-anonymity could be applied for both dimensions: *spatial* and *temporal* (Gedik, & Liu, 2005, 2008). In this approach the user has the ability to define upper and lower limits for both area size and time periods.

2.1.4 Obfuscation

The idea of obfuscation is to reduce location precision of the user. The traditional approach is sending a circle area rather than the true accurate position. (Duckham, & Kulik, 2005) apply this concept to the road networks. Again, for this approach *spatial* obfuscation can be improved by adding *temporal* obfuscation (Gruteser, & Grunwald, 2003).

2.1.5 Coordinate Transformation

User's true position can be protected by applying some simple transformations (shifting, rotating... etc) on the coordinates. These transformations should be known at the receiver side (Gutscher, 2006).

2.1.6 Cryptography

Encryption concept is used in this approach to protect the true position of the user. Symmetric encryption techniques are usually used. To deal with the problem of untrusted TTP, (Marias, Delakouridis, Kazatzopoulos, & Georgiadis, 2005) proposed a secret sharing based approach, where the shares are distributed among many untrusted TTP. Thus, to retrieve the true position, you need all the shares. The dilemma here is whether we can perform location-based queries using these shares or not.

2.1.7 Position Sharing

In order to overcome the problem of applying LBS queries on sub shares, (Dürr, Skvortsov, & Rothermel, 2011) proposed the concept of shares with strictly limited precision. Each share has its own role to accurately determine the true position of the user. Even if one or more shares are not accessible, we can apply LBS queries on the remaining shares.

2.2 Privacy Preferences Settings Approaches

In Location-Based services, Users have the ability to configure their privacy preferences settings manually. But there are some usability issues with this approach. In (Sadeh et al., 2009) study, they found that many users face difficulties in order to set their preferences in the proper way. On the other hand, users' decisions about their location sharing preferences are dependent on the current context (Anthonyet al., 2007). If contexts are included in the manual settings, it will be more complex.

Machine Learning approaches were proposed to predict user's decisions about their location sharing preferences. Privacy wizard for the social networks was proposed by (Fang, & LeFevre, 2010). They built an adaptive binary classifier guided by some answers for users' privacy preferences. Random Forest as a machine learning method was applied by (Sadeh et al., 2009). According to their results, machine learning methods perform better than user-defined rules. Machine learning methods need to learn from the previous decisions of the same user to well predict the decisions for new contexts. This is why they can't perform well during cold-start periods.

To deal with the issue of cold-start periods, other approaches were proposed. These approaches depend on the decisions from other users rather than the new user. The idea behind these approaches is that if the behavior of the users is similar during past decisions, then they may behave the same during future ones. Semantic categories based method was proposed by (Toch, 2014). Semantic categorization for locations is done first. Then the decisions are collected from users with the same semantic category. (Jin, Saldamli, Chow, & Knijnenburg, 2013) method depends on the locations activities in order to select decisions from similar users.

According to the value of decisions whether it is discrete or continuous, used approaches could be classified into *discrete opinion dynamics models* and *continuous opinion dynamics models* (Shang, 2014). In this work we are concentrating on *User-User collaborative filtering (CF)* which is considering as *discrete opinion dynamics model.* In *User-User CF* the underlying opinion interaction mechanism is not explicitly involved in the matrix representation. Understanding the underlying mechanism of opinion dynamics could be achieved in more details using *continuous opinion dynamics models* (Shang, 2013, 2014) which are considered as a possible extension for our work.

2.3 User-User Collaborative Filtering

Recommender for location privacy preferences based on *User-User collaborative filtering (CF)* was proposed by (Desrosiers, & Karypis, 2011). Decisions are collected from the users with the highest similarities. This method consists of three stages:

2.3.1 Generating Rating Vectors

In CF recommender models, the decision have the form of <user; context, decision>. Contexts here are combination of locations and times. Locations and times were divided into location group L and time group T respectively. Thus the total number of contexts is:

$$C=L*T \tag{1}$$

Decisions are either positive or negative. Positive decision has the (5) value and means that the decision of the user is to share the location. Negative decision has the (1) value and means that the decision of the user is not to share the location.

Table 1. shows a sample of rating vectors for 5 users and 4 contexts. Each row represents the rating vector for one user. Some users may have NULL value for some entries. This mean that this user does not made a decision for such context. Contexts are combination of location and time, for example C1 could be (*Library, Morning*).

Table 1. Rating vectors sample for 5 users and 4 contexts

user	C_1	C_2	C_3	C_4	C_4
u_1	1	5	1	5	1
u_2	5	5	1	NULL	1
u_3	1	5	1	5	**???**
u_4	1	5	1	5	5
u_5	1	5	1	5	5

2.3.2 Calculating User Similarity

The Second stage is to find the users with highest similarity with the needed user. Those users should have a decision for the intended context, and have decisions for some contexts that the intended user also has decisions. Cosine similarity was used. But before calculating the similarity, each vector is normalized first. Normalization step is needed because some users tend to be positive for their sharing while others tend to be negative. Normalization is done by subtracting the mean of decisions for each user from his decisions.

2.3.3 Predicting Privacy Preferences

After finding the similar users the predicted decision for user *u* and item *i* is as the following (Desrosiers, & Karypis, 2011):

$$\hat{r}_{ui} = \bar{r}_u + \frac{\sum_{v \in N_i(u)} \; w_{uv}(r_{vi} - \bar{r}_v)}{\sum_{v \in N_i(u)} \; |w_{uv}|} \tag{2}$$

Where *Ni(u)* is the set of similar users, *v* is any user from this set, r_v is the mean rating for user *v*, w_{uv} is the similarity weight between users *u* and *v*. Then the final decision is a value between 1 and 5, and they use 3 as threshold as the following (Desrosiers, & Karypis, 2011):

$$R_{u,i} = \begin{cases} negative & \text{if } \hat{r}_{u,i} \leq \theta \\ positive & \text{if } \hat{r}_{u,i} > \theta \end{cases} \tag{3}$$

For the example in Table.I we can see that it is clear to show that the decision will be 5, which is positive.

2.3.4 Obfuscating Rating Vectors

This step is additional and it was added to achieve the privacy for the users. Fake ratings were added in the *NULL* entries. If m_t denotes the number of rated items for user *u*, α is the noise factor, then $m_{max}=\alpha m_t$ is the upper limit of added fake items. The number of added fake items m_f is generated between 0 and m_{max} and belongs to uniform distribution. Finally they apply their algorithm to real world dataset. In general, the results were comparable to the performance of Machine Learning based models and better in the cold start period.

(Xie, Knijnenburg, & Jin, 2014) propose the idea of integrating both *User-User collaborative filtering* and *Item-Item collaborative filtering*.

3. Proposed Methods

The main purpose of this paper is to improve the work of (Zhao et al., 2014). The improvements were achieved through the following aspects:

- ✓ Discuss the positive and negative aspects of using both Collaborative based methods and Machine Learning based methods
- ✓ Enhance the prediction by Applying Improved Machine Learning method
- ✓ Combine both Collaborative based methods and Machine Learning based methods in one integrated model

3.1 Collaborative Based Methods and Machine Learning Based Methods

For location privacy preferences, both of these methods try to predict preferences according to previous data. Machine learning methods depend on the previous data for the intended user. On the other hand, collaborative based methods depend on the data collected from other users. According to this point, collaborative based methods perform better during the “cold-start” period. The performance of machine learning methods with sufficient training data is better.

Machine learning methods are sensitive to the change of the data such as data obfuscating. (Zhao et al., 2014) have successfully applied data obfuscating. Although they were considering semi-honest TTP, the machine which generates rating vectors still need the data as is before obfuscating. Thus the privacy issue still stands for at least this machine.

Machine learning methods have the ability to replace the classifier in a plug-out, plug-in way. Thus we can test the impact of different classifiers with less effort.

An important point is the impact of adding new features for both methods. In machine learning methods, adding new features is considerably easier to achieve. Machine learning methods are designed to deal with high dimensional features. In collaborative based methods, the contexts are generated from the features. Thus adding new feature will increase the size of rating vectors with the multiplication of the distinct elements of that feature. In addition, features values should be grouped and rounded in order to generate the contexts from the features easily. Furthermore, sometimes it is not easy to group the values of the feature in one unified group set for all users.

To clarify this point, let's consider the dataset and the work accomplished by (Zhao et al., 2014). They used two features: location and time. Locations and times were divided into location group L and time group T respectively as the following:

$$L = \{\text{Food and Drink; Leisure; Retail; Residential; Academic; Library}\}$$

$$T = \{\text{Morning; Noon; Afternoon; Evening; Night}\}$$

The contexts were a combination of locations and times. Thus the total number of contexts is $6 \times 5 = 30$. Let's consider the impact of adding the Friend List ID as a new feature with a group size of 5. The new total number of contexts is $6 \times 5 \times 5 = 150$, which is five times the original size. On the other hand, grouping the values of the Friend List ID in one unified group set for all users is difficult. Even if we consider the type of Friend List rather than the ID, the problem will remain since it could be customized and may be deferent from user to user.

3.2 Applying Improved Machine Learning Method

The second direction in this paper is to enhance the prediction of location privacy preferences by applying improved machine learning methods. This improvement is done by adding new significant features. As we mentioned before it is easy to add new features in machine learning method.

(Zhao et al., 2014) work depends on just two features: Location and time. Although these features are important for the decision of sharing, the Friend List feature and the size feature of that Friend List is more important.

In this paper we will improve the prediction performance of location privacy preferences by adding these features.

Random forests (Breiman, 2001) were applied as machine learning method. It constructs decision trees using the features in order to predict the decision. It uses ensemble learning method which means that it will constructs more than one decision trees and the mode of the results will be selected.

3.3 Collaborative Based Methods and Machine Learning Based Methods as One Integrated Model

The Third direction is to combine both Collaborative based methods and Machine Learning based methods in one integrated model. Here we will use user-user CF as collaborative based method and Random forests as machine learning method. Figure.1 shows the flowchart for the proposed integrated model. The concept is that during training phase, both of the two methods will be evaluated using cross validation technique. The training data itself will be divided into k sub-sets. Then k-1 sets will train the predictor and one set will be used as testing. This step will be repeated for each set. Finally, the mean of the accuracy for the k tested sets will be used as evaluation value. According to this value we will select the corresponding learned method.

4. Results and Discussion

In order to illustrate the proposed enhancement, the same real world dataset used in the study of (Zhao et al., 2014) was used here.

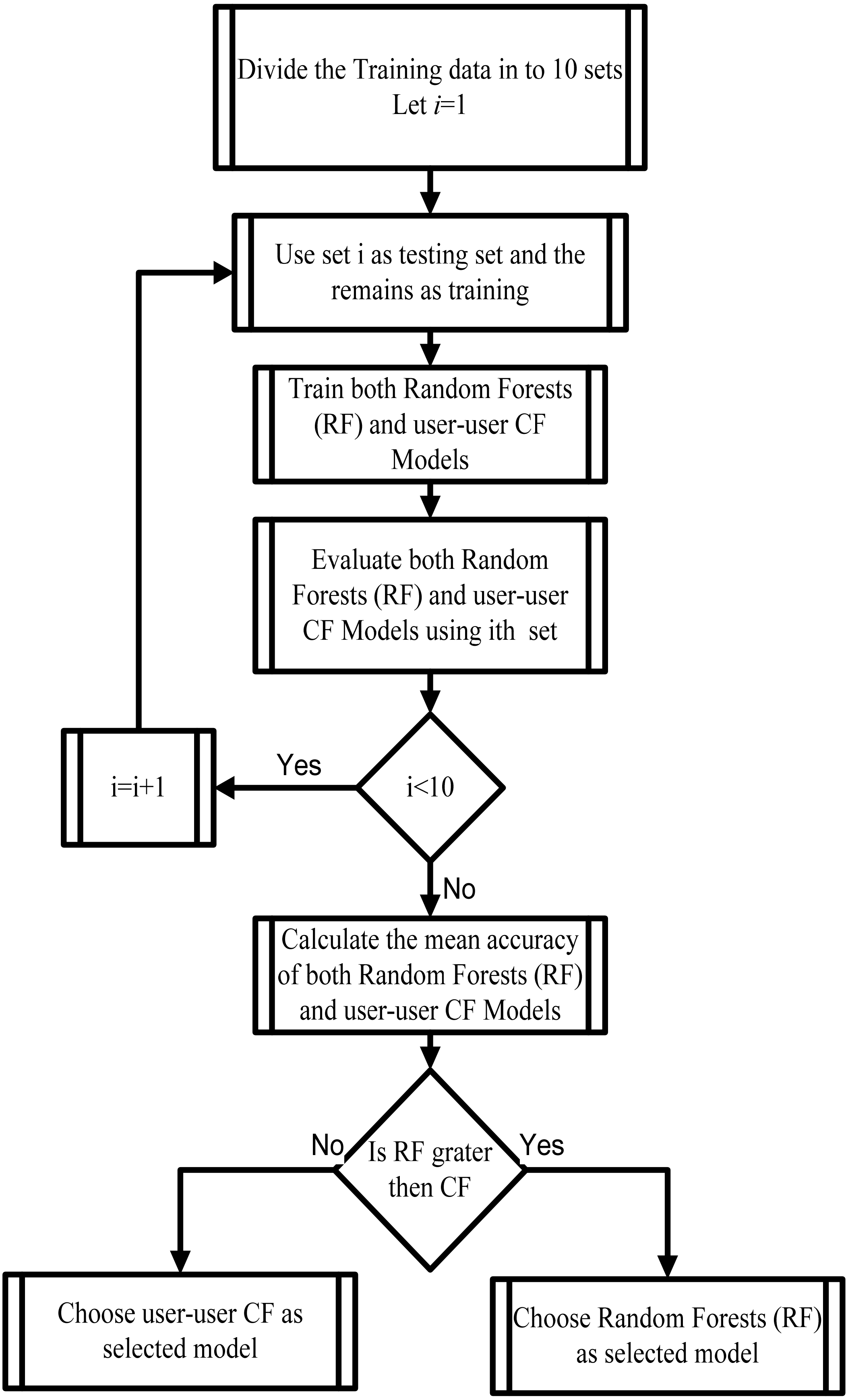
Divide the Training data in to 10 sets
Let *i*=1
Use set i as testing set and the remains as training
Train both Random Forests (RF) and user-user CF Models
Evaluate both Random Forests (RF) and user-user CF Models using ith set
i<10
Yes
i=i+1
No
Calculate the mean accuracy of both Random Forests (RF) and user-user CF Models
Is RF grater then CF
No
Yes
Choose user-user CF as selected model
Choose Random Forests (RF) as selected model

Figure 1. Proposed integrated model

4.1 Used Dataset

In this study LocShare dataset (Parris, & Abdesslem, 2011) was used. This data set represents real location privacy preference data. The experiment was conducted in both St Andrews and London. Two runs for each place. In order to compare our results with (Zhao et al., 2014) work, we just include St Andrews data. The data from each run of the experiment are contained in five csv files:

- ✓ Users: the users participating in the experiment;
- ✓ Acc: the accelerometer data;
- ✓ Events: the events that took place involving the user, and responses to questions regarding those events;
- ✓ Encounters: the encounters between users;
- ✓ FriendsLists: details regarding each participant's friendlists.

The most important ones for our study are *Events* and *FriendsLists*. Events file contains the events readings collected during the experiment. It has eight different features:

- ✓ User ID
- ✓ Type of event
- ✓ Time question was asked
- ✓ Time the question was answered
- ✓ The place category
- ✓ The List ID if the user wants to share with.
- ✓ Response
- ✓ Co-presence with friends

The data consists of 18,153 records. Records without responses and without place categories were ignored. The final number was 3,878 records. The time was grouped as the following {7:00-12:00; 12:00-14:00; 14:00-17:00; 17:00-21:00; 21:00-7:00} for the groups {Morning; Noon; Afternoon; Evening; Night} respectively. The size of the List ID was retrieved from *FriendsLists* file.

4.2 Used Metrics

In our experiment the results were evaluated using:

Accuracy: This represents the percentage of correctly predicted decisions.

Privacy leaks: is another important metric which represents the percentage of the prediction overexposed of users' location information. This is calculated using the negative decisions that are predicted as positive ones. Privacy leak is also called False Positive (FP) in the terminology of machine learning.

4.3 Evaluation Experiment

In order to conduct the experiment MATLAB R2014a (The MathWorks, 2014) was used as programing tool. The dataset was processed and imported to MATLAB. WEKA 3.6.12 (Hall et al., 2009) as machine learning classifiers package was called from MATLAB to use Random Forests. This link between MATLAB and WEKA was accomplished using (Dunham, 2008) code. User-User CF was re-implemented as it was described in (Zhao et al., 2014).

For both methods, the data for each user was divided into 10 sets randomly in order to apply the cross validation. The first set is used as testing, while the rest of the sets were merged to provide the training data. Using this training data, each method was learned. Then, for the test set, the true decisions were used to evaluate the performance of each method .This step (training and testing) was repeated 10 times, one time for each set. And the whole experiment was repeated 10 times. Both of the Accuracy and Leak metrics were calculated. And the mean value over all experiment was considered. Figure 2 shows the flowchart for the evaluation experiment.

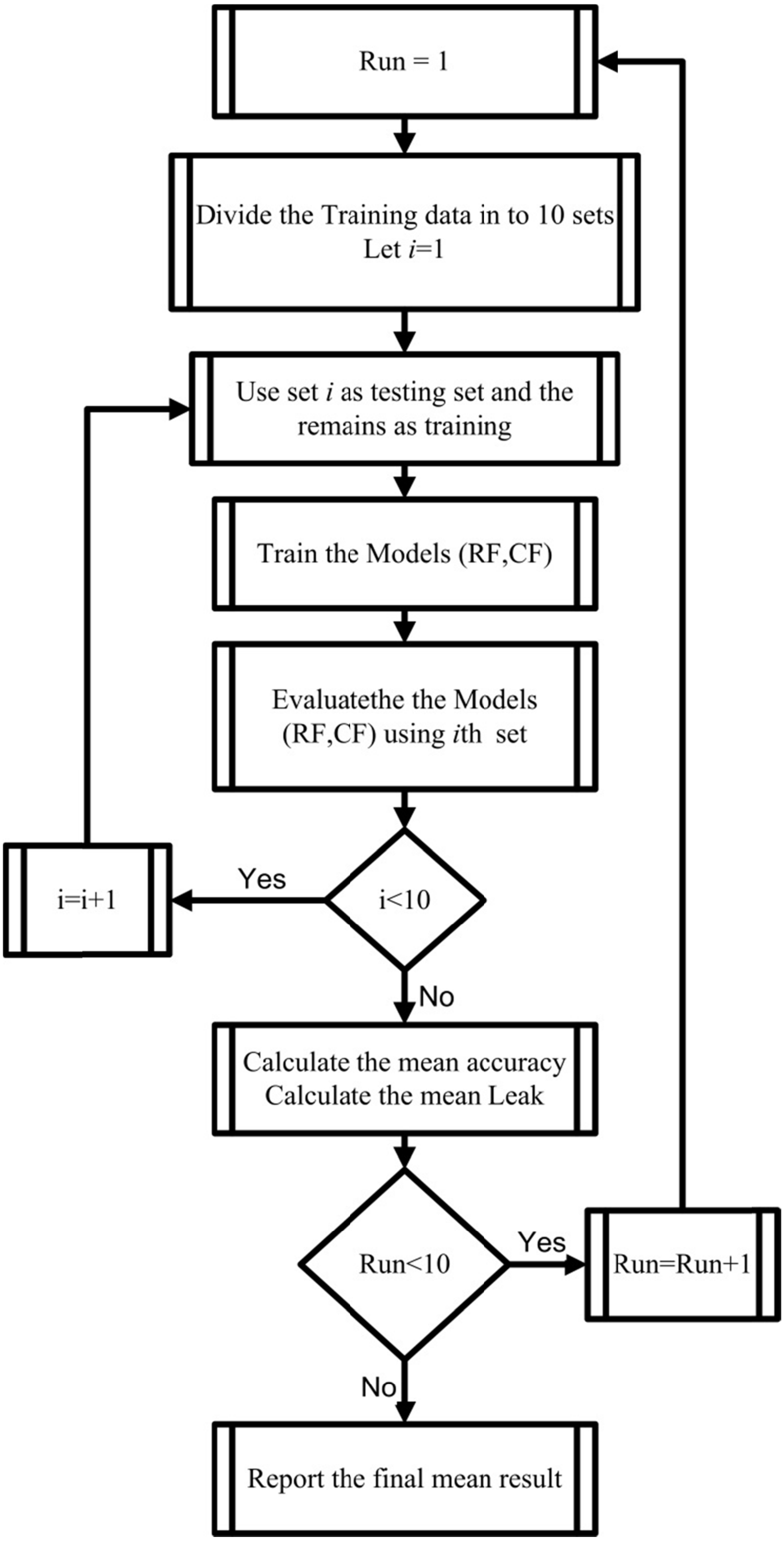
Run = 1
Divide the Training data in to 10 sets
Let *i*=1
Use set *i* as testing set and the remains as training
Train the Models (RF,CF)
Evaluatethe the Models (RF,CF) using *i*th set
i<10
Yes
i=i+1
No
Calculate the mean accuracy
Calculate the mean Leak
Run<10
Yes
Run=Run+1
No
Report the final mean result

Figure 2. Evaluation Experiment

For *User-User CF* the training data was combined with the data from other users. Thus *User-User CF* can collect the decisions from other users. On the other hand, for *Random Forests,* the data from other users are not used since we depend on the history of the intended user only.

Figure 3 illustrates the improved results using *Random Forests* –the green circle- with accuracy (84.82%) and privacy leak of (7.65%). Clearly this result outperforms the *User-User CF* –the red rectangle- where the accuracy was (72.5%) and the leak was (13.4%). *User-User CF* –the black rectangle-results reported by (Zhao et al., 2014) was (73.11%) for the accuracy and (11.75%) for the leak. The deference here may be due to some tiny implementation details. Finally we can see the effect of adding new features from the deference in the between *Random Forests* results achieved here. And *Random Forests* –the blue circle- results reported in (Zhao et al., 2014).

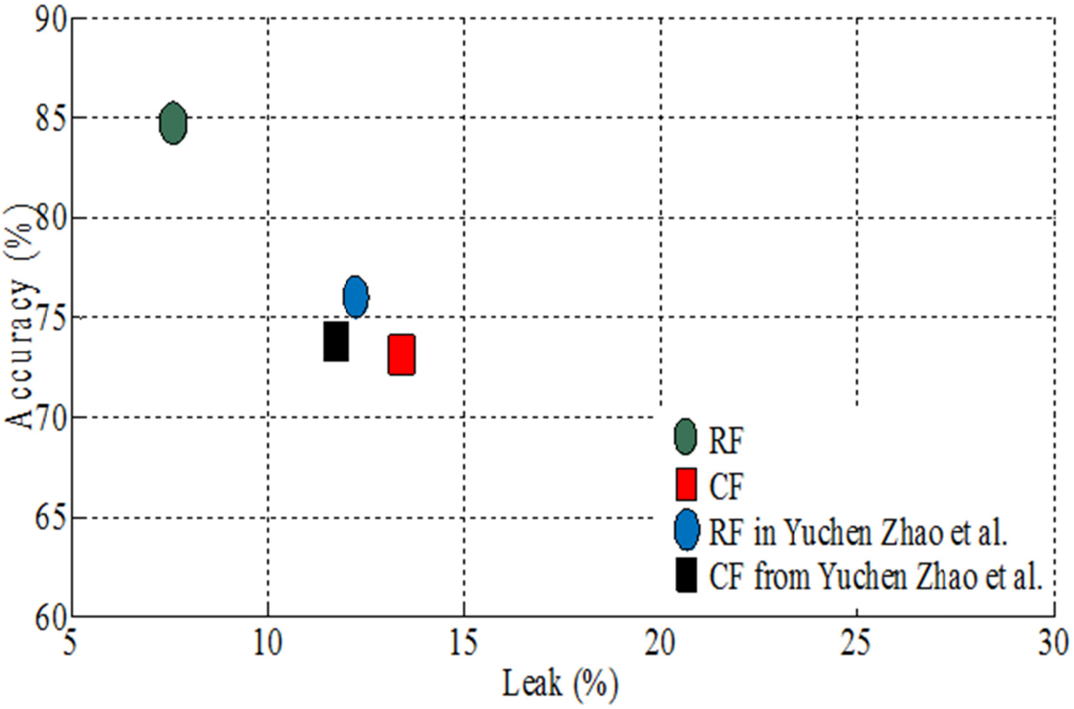

Figure 3. Comparing RF and CF

Figure 4 shows the impact of obfuscating rating vectors on the accuracy. The red line is the baseline with no obfuscating. The blue line represents the effect of adding fake rating into User-User CF.

Alpha here is the noise factor. From the equation $m_{max}=\alpha m$, we can see that the impact was a tiny decrease in the accuracy (around 3.5%). obfuscated rating vectors scaled well with the increase of the noise. This is due to the fact that alpha has an impact of the upper limit of randomly selected value m_f between 0 and m_{max}. One the other hand the added fake ratings are divided into two equal sets: positives and negatives. Thus the mean of these fake rating is 0 which has no effect on the calculations.

The third experiment introduces the concept of combining both *User-User CF* and *Random Forests*. The idea is to take the advantages of both machine learning approaches and collaborative approaches. In specific, gaining high accuracy during the "cold-start" period and when there is sufficient data to train the system.

The concept here is that during training phase, both of the two methods will be evaluated using 10-k cross validation technique. The training data itself will be divided into 10 sub-sets. One set will be used for testing, while the remaining will be used for training.

This step will be repeated for each set and the estimated accuracy will be calculated. Finally, the mean of the estimated accuracies will be used as evaluation value. According to this value, we will select the corresponding learned method.

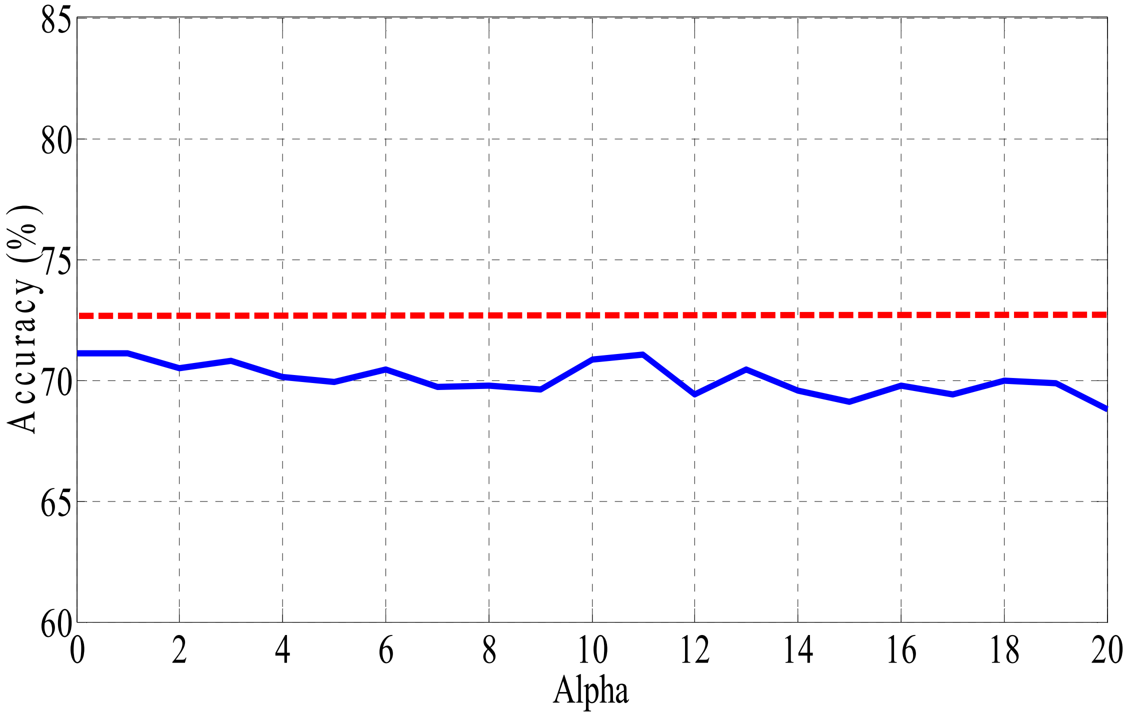

Figure 4. CF with obfuscating rating vectors

In Figure 5 we study the impact of reducing the training data. We can see that *Random Forests* outperforms *User-User CF* during the whole interval. This is due to the high difference between them in accuracy (around 12.5%). Thus even during the cold start period, *Random Forests* still outperforms *User-User CF*. We can also see that *User-User CF* scales better than *Random Forests*. This is normal due to the fact that *User-User CF* takes the advantage of the available data for the intended user, and the rest of the users.

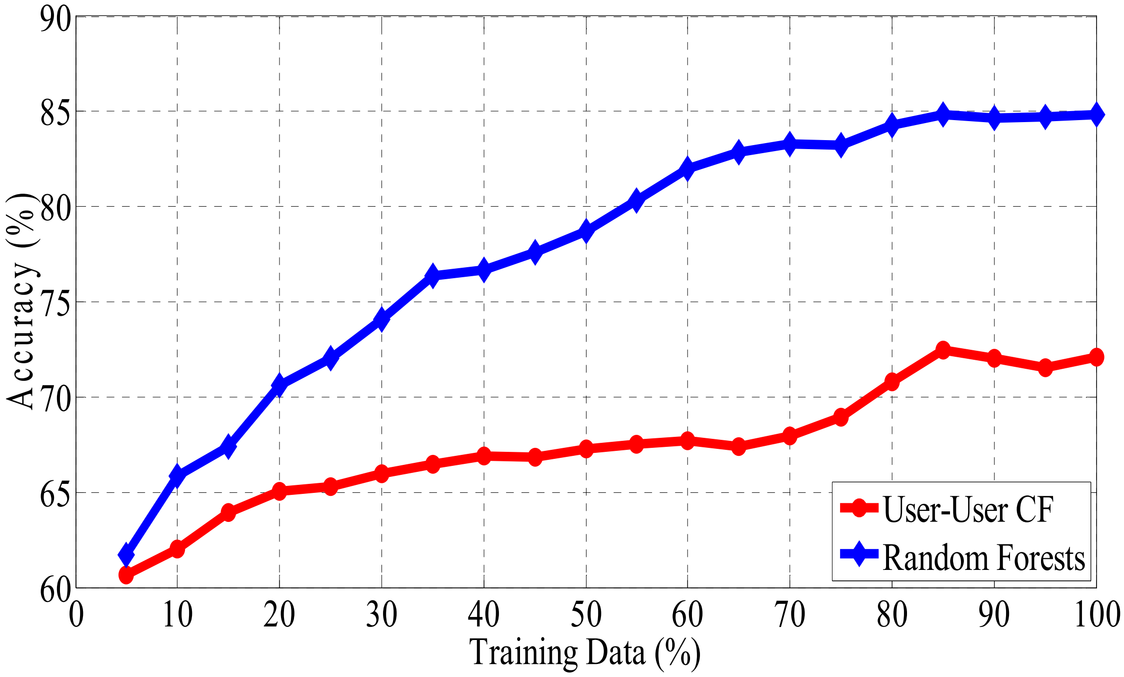

Figure 5. The impact of reducing the training data

According to this observation, the performance of the integrated model will behave similarly to *Random Forest* performance.

5. Conclusion

This paper starts by illustrating some positive and negative aspects of using both Collaborative based methods and Machine Learning based methods. We show that Collaborative based methods have the theoretical advantages of good performance during "cold-start" period. It learns from other users' data allowing us to apply the obfuscating by adding some fake decisions. The main drawback of these methods is the complexity of dealing with new features. On the other and, Machine Learning based methods have the advantage of better performance. In general, adding new features is very easy, and we can replace the classifiers easily. But we mentioned also that it suffers from the poor performance during "cold-start" period, the model can learn from the history of the intended user only, and adding fake decisions will affect the performance.

After that, we have enhanced the prediction of location privacy preferences by applying improved *Random Forests as* machine learning method. We show that the results out performs *User-User CF* with accuracy of (84.82%) and privacy leak of (7.65%).

Finally, we proposed the concept of combining both Collaborative based methods and Machine Learning based methods in one integrated model to utilize the advantages of both of them. But we observed that *Random Forests* outperforms for the whole period. Thus, the integrated method will behave the same as *Random Forest* methods.

References

Anthony, D., Henderson, T., & Kotz, D. (2007). Privacy in location-aware computing environments. *IEEE Pervasive Computing,* (4), 64-72. http://dx.doi.org/10.1109/MPRV.2007.83

Breiman, L. (2001). Random forests. *Machine learning, 45*(1), 5-32. http://dx.doi.org/10.1023/A:1010933404324

Desrosiers, C., & Karypis, G. (2011). *A comprehensive survey of neighborhood-based recommendation methods.* In Recommender systems handbook (pp. 107-144). *Springer US*. http://dx.doi.org/10.1007/978-0-387-85820-3_4

Duckham, M., & Kulik, L. (2005). A formal model of obfuscation and negotiation for location privacy. In Pervasive computing (pp. 152-170). *Springer Berlin Heidelberg.* http://dx.doi.org/10.1007/11428572_10

Dunham, M. (2008). *Matlab Weka Interface.* Retrieved from http://www.mathworks.com/matlabcentral/fileexchange/21204-matlab-weka-interface/content/matlab2weka.m

Dürr, F., Skvortsov, P., & Rothermel, K. (2011). *Position sharing for location privacy in non-trusted systems*. In Pervasive Computing and Communications (PerCom), 2011 IEEE International Conference on (pp. 189-196). IEEE. http://dx.doi.org/10.1109/PERCOM.2011.5767584

Fang, L., & LeFevre, K. (2010, April). Privacy wizards for social networking sites. In Proceedings of the 19th international conference on World wide web (pp. 351-360). *ACM.* http://dx.doi.org/10.1145/1772690.1772727

Gedik, B., & Liu, L. (2005). *Location privacy in mobile systems: A personalized anonymization model.* In Distributed Computing Systems (pp. 620-629). ICDCS 2005. Proceedings. 25th IEEE International Conference on IEEE. http://dx.doi.org/10.1109/ICDCS.2005.48

Gedik, B., & Liu, L. (2008). Protecting location privacy with personalized k-anonymity: Architecture and algorithms. *Mobile Computing, IEEE Transactions on, 7*(1), 1-18. http://dx.doi.org/10.1109/TMC.2007.1062

Gruteser, M., & Grunwald, D. (2003). Anonymous usage of location-based services through spatial and temporal cloaking. In Proceedings of the 1st international conference on Mobile systems. *Applications and Services* (pp. 31-42). ACM. http://dx.doi.org/10.1145/1066116.1189037

Gutscher, A. (2006). *Coordinate transformation-a solution for the privacy problem of location based services?* In Parallel and Distributed Processing Symposium, 2006. IPDPS 2006. 20th International IEEE. http://dx.doi.org/10.1109/IPDPS.2006.1639681

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. *ACM SIGKDD Explorations Newsletter, 11*(1), 10-18. http://dx.doi.org/10.1145/1656274.1656278

Jin, H., Saldamli, G., Chow, R., & Knijnenburg, B. P. (2013). *Recommendations-based location privacy control*. In Pervasive Computing and Communications Workshops (PERCOM Workshops), 2013 IEEE International Conference on (pp. 401-404). IEEE. http://dx.doi.org/10.1109/PerComW.2013.6529526

Krumm, J. (2007). *Inference attacks on location tracks*. In Pervasive Computing (pp. 127-143). Springer Berlin Heidelberg. http://dx.doi.org/10.1007/978-3-540-72037-9_8

Marias, G. F., Delakouridis, C., Kazatzopoulos, L., & Georgiadis, P. (2005). *Location privacy through secret sharing techniques*. In World of Wireless Mobile and Multimedia Networks, 2005. WoWMoM 2005. Sixth IEEE International Symposium on a (pp. 614-620). IEEE. http://dx.doi.org/10.1109/WOWMOM.2005.60

Mokbel, M. F., Chow, C. Y., & Aref, W. G. (2006). *The new Casper: query processing for location services without compromising privacy*. In Proceedings of the 32nd international conference on Very large data bases (pp. 763-774). VLDB Endowment.

Palanisamy, B., & Liu, L. (2011). *Mobimix: Protecting location privacy with mix-zones over road networks*. In Data Engineering (ICDE), 2011 IEEE 27th International Conference on (pp. 494-505). IEEE. http://dx.doi.org/10.1109/ICDE.2011.5767898

Parris, I., & Abdesslem, F. B. (2011). *CRAWDAD data set st_andrews/locshare* (v. 2011-10-12)

Sadeh, N., Hong, J., Cranor, L., Fette, I., Kelley, P., Prabaker, M., & Rao, J. (2009). Understanding and capturing people's privacy policies in a mobile social networking application. *Personal and Ubiquitous Computing, 13*(6), 401-412. http://dx.doi.org/10.1007/s00779-008-0214-3

Shang, Y. (2013). Deffuant model with general opinion distributions: First impression and critical confidence bound. *Complexity, 19*(2), 38-49. http://dx.doi.org/10.1002/cplx.21465

Shang, Y. (2014). An agent based model for opinion dynamics with random confidence threshold. *Communications in Nonlinear Science and Numerical Simulation, 19*(10), 3766-3777. http://dx.doi.org/10.1016/j.cnsns.2014.03.033

Shankar, P., Ganapathy, V., & Iftode, L. (2009). *Privately querying location-based services with SybilQuery*. In Proceedings of the 11th international conference on Ubiquitous computing (pp. 31-40). ACM. http://dx.doi.org/10.1145/1620545.1620550

The MathWorks. (2014). *MATLAB Toolbox Release R2014a*, Natick, Massachusetts, United States.

Toch, E. (2014). Crowdsourcing privacy preferences in context-aware applications. *Personal and ubiquitous computing, 18*(1), 129-141. http://dx.doi.org/10.1007/s00779-012-0632-0

Wernke, M., Skvortsov, P., Dürr, F., & Rothermel, K. (2014). A classification of location privacy attacks and approaches. *Personal and Ubiquitous Computing, 18*(1), 163-175. http://dx.doi.org/10.1007/s00779-012-0633-z

Xie, J., Knijnenburg, B. P., & Jin, H. (2014). *Location sharing privacy preference: Analysis and personalized recommendation*. In Proceedings of the 19th international conference on Intelligent User Interfaces (pp. 189-198). ACM. http://dx.doi.org/10.1145/2557500.2557504

Zhao, Y., Ye, J., & Henderson, T. (2014). *Privacy-aware location privacy preference recommendations*. In Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (pp. 120-129). ICST (Institute for Computer Sciences, Social-Informatics and

Telecommunications Engineering). http://dx.doi.org/10.4108/icst.mobiquitous.2014.258017

Copyrights