Item Recommendation Using Collaborative Filtering in Mobile Social Games: A Case Study

Zhaojie Tao*, Ming Cheung†, James She‡
HKUST-NIE Social Media Lab
Hong Kong University of Science & Technology, Hong Kong
*ztao@connect.ust.hk, †cpming@ust.hk, ‡eejames@ust.hk

Ringo Lam§
nxTomo Games Limited & Tsinghua University
§ringo.lam@nextmedia.com

Abstract—This paper evaluates the performance of collaborative filtering in mobile social game. The evaluation involves both user-based and item-based collaborative filtering on game items for in-app purchases, and including 4 different social information available in the game. Based on the operational data from a mobile social game, Barcode Footballer, with more than 100k users and 50k purchasing history, it is concluded that both user-based and item-based collaborative filtering have much higher precision than random recommendation, while user-based approach with friendship as similar relationship has better performance than original approach. This paper also proposes a hybrid method to improve the performance of user-based friendship approach. The results can be applied to mobile social games to recommend highly needed items to users so that the monetization can be enhanced.

Keywords—*Recommendation; collaborative filtering; in-app purchases; game items; social mobile game*

I. INTRODUCTION

Mobile social games are based on social networks and mobile technologies, which integrate social features as gameplay elements and suitable to play on mobile devices. Mobile social game industry has grown rapidly and altered the game industry deeply. According to a research by SuperData, there were 77.9 million social network players in US by the end of 2012 and the market is 6.2 billion USD, which is expected to reach 8.64 billion USD in 2014. Since mobile social game is a new developing industry, the ways to design a successful mobile social game is still a challenge for all the game companies. One of the analytic frameworks in game development is ARM (Acquisition, Retention, Monetization). Acquisition is to attract users to play the game, Retention focus to keep users in the game, and Monetization is how operators gain revenue. One of the common approach of monetization is using in-app purchase. According to research by Distimo, almost all revenue in free apps are from in-app purchases, which indicates the importance of in-app purchases. In a mobile monetization report released by Swrve, that half of mobile game money comes from 0.15% of players. As a result, the ways to monetize more paid users is a prefered challenge for game companies, one possible solution is using recommendation.

At present, recommender systems can be classified into three types, namely content-based recommendations, collaborative filtering recommendations, and hybrid recommendations [1]. In content-based approach, items of the user is estimated by the similarity among items. In collaborative filtering approach, the utility of items for a user is predicted based on

the items that rated by other users. The hybrid approach is a combination of content-based approach and collaborative filtering approach. This paper focuses on the collaborative filtering approach by investigating a mobile social game, called Barcode Footballer.

Barcode Footballer is a football game, it is currently launched by nxTomo, which is a Hong Kong based company. In this game, each user can create and manage a football team, shown in Fig. 1.(a) As a mobile social game, it includes 4 social elements, i.e., Friendship, Location, League, and Level. The game have more than 1 million friendships, users can find out new friends in Friends Searching interface. Users can change the football team with better footballers and can join matches and leagues in order to rise to higher levels and enter better leagues. Besides, the game includes over 100 locations in Hong Kong, which indicates the places that users registered in the game. In addition to the normal matches with default players, users can also join the matches with other users, which are based on the users location, so that users can play the game with other nearby users. As a free download game, it was designed with in-app purchase function. Users can purchase items from the item store, shown in Fig. 1.(b), the items in this game can be classified into three types, namely basic items, football shirts and team logos. Items can be purchased by using game point or club fund, which can be purchased in app or raise in game respectively. Since in-app purchase can generate great revenue for game companies, how to encourage users purchase virtual goods is one of the most important question in monetization. This paper focuses on recommending game items, which is a popular in-app purchasing virtual good in mobile social games. It investigates the user data from Barcode Footballer and tries to compare the performance of collaborative filtering with different conditions.

To the best of our knowledge, there is no previous investigation on real datasets to evaluate the performance of collaborative filtering on game item for in-app purchases. The contributions of this paper can be summarized as follows:

1) Investigate a social mobile game with 100k+ users, Barcode Footballer;
2) Evaluate the performance of collaborative filtering in mobile social game in-app purchases;
3) Propose a modified user-based collaborative filtering as well as a hybrid method for item recommendation.

The paper is organized as follows: Section II describes the previous works, Section III introduced the methodology in game item recommendation, Section IV describe the details

IEEE computer society

Fig. 1. (a) Team Management Interface of Barcode Footballer. (b) Item Purchase Interface of Barcode Footballer.

Fig. 2. Find out Relevant Items in Item-based Scenario.

of dataset, and the evaluation results with analysis, Section V proposes a combined method, Section VI is the conclusion.

II. RELATED WORKS

Using an in-app purchase model in mobile apps is a more effective method of converting casual users into paying customers. It is a different approach from the traditional payment where users pay and download, which the users may be disappointed by the experience after purchasing [2, 3]. In-app purchasing provides a recurring revenue stream for game developers, this enhances the importance of analyzing the in-app purchasing pattern and find out how to increase the monetization [4]. In [5], the authors analyze the in-app purchasing patterns in a social game called PuppyRed. The items can be purchased by coin (real money) or bean (obtained in game), which is similar with Barcode Footballer. The authors classified the game items into 3 types, which distinguish items according to use items on the role, space and the gameplay. The result shows that high spenders like to purchase items by coins for the role and space, while low spenders like to purchase items by bean for the gameplay. In [6], the authors analyze the users purchasing pattern and suggest how to distinguishing money spenders and non-spenders.

In addition to purchasing pattern, recommendation is also an important method to enhance revenue [7]. As previous works show, collaborative filtering recommendation can be classified into two types, namely user-based approach and item-based approach. In [8], both of this two approaches are
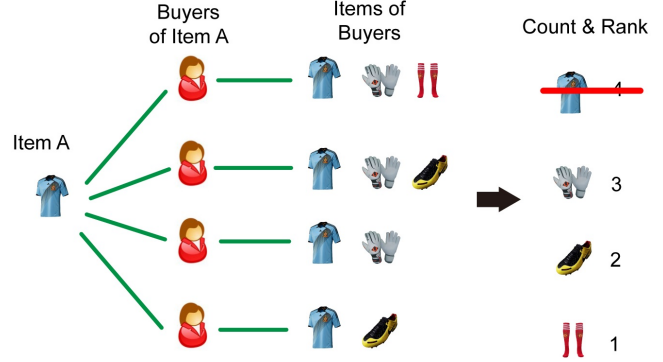
evaluated in mobile game recommendation. The result shows that item-based collaborative filtering with combined explicit and implicit data has the best performance (largest AUC in ROC curve), and in item-based approach, using Pearson Correlation calculate the similarity has better performance, while in user-based approach, Euclidean Distance is better. In [9], they discuss how recommendation algorithm could be evaluated in order to select the best algorithms from a set of candidates.

III. METHODOLOGY

A. Item-based Collaborative Filtering

In item-based scenario, for each item, recommendation is made according to the items of the buyers. For instance, for an item *A*, collect all the items of item *A* buyers, shown in Fig. 2. Firstly count the repeated times of each item, and then sort all the items, and choose the top-*k* items as the most related items of item *A*. In other words, it reveals that the users who have purchased item A also purchased top*k* items. After doing this for all the items, a result that include most connected items of each item can be get.

Algorithm 1 Find out most related items for an item

Input: Item buyers' purchased items, $\{Pocket_u\}$;
Output: Item list, $\{Item_i\}$; Times of items be purchased, $\{Count_i\}$;
1: **for** each $Pocket_x$ of user x in $\{Pocket_u\}$ **do**
2:     **for** each $Item_y$ in $Pocket_x$ **do**
3:         **if** $Item_y$ in $\{Item_i\}$ **then**
4:             $Count_x$++;
5:         **else**
6:             Add $Item_y$ to $\{Item_i\}$;
7:         **end if**
8:     **end for**
9: **end for**
10: Sort $\{Item_i\}$ according to $\{Count_i\}$;
11: **return** $\{Item_i\}$;

The Algorithm 1 describes how to find out what other items do the users most likely to buy when they purchase an item. After using the Algorithm 1 for every items of the game, each item can get a list of most related items.

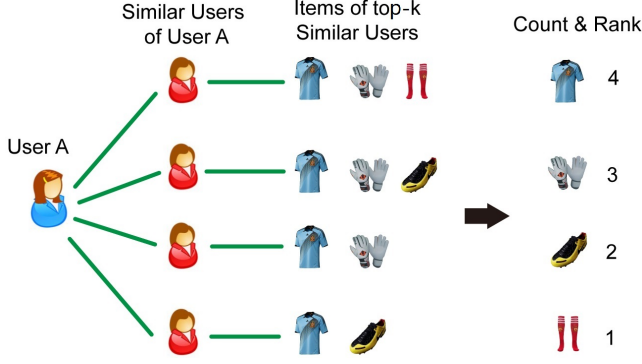To make the recommendation, the data of purchasing

Fig. 3. Find out Relevant Items in User-based Scenario.

Fig. 4. Distribution of users based on how many items they purchased.

record is analyzed. If a user buy an item, a list of most related items of that item can be listed quickly, and then choose top-*N* items from the list, where variable *N* is the number of recommendations.

B. Original User-based Collaborative Filtering

In user-based scenario, recommendation will be made based on similar users, in terms of users' purchasing records. Firstly, setting up item rating tables, where the rating is obtained from purchasing records. Secondly, using the rating table to calculate the similarity between each two users in terms of purchased items. Finlly, recommend the top-*N* purchased items from top-*k* similar users, shown in Fig. 3. The algorithm of user-based collaborative filtering is shown in Algorithm 2, for each user, a list of items can be generated, and then choose top-*N* items from the list according to the recommendation amount requirement.

Algorithm 2 Make user-based collborative filtering

Input: Similar users list, $\{U\}$; Item buyers' purchased items, $\{Pocket_u\}$;

Output: Item list, $\{Item_i\}$; Times of items be purchased, $\{Count_i\}$;

 1: **for** each user x in $\{U\}$ **do**
 2:     **for** each $Item_y$ in $Pocket_x$ **do**
 3:         **if** $Item_y$ in $\{Item_i\}$ **then**
 4:            $Count_x$++;
 5:         **else**
 6:            Add $Item_y$ to $\{Item_i\}$;
 7:         **end if**
 8:     **end for**
 9: **end for**
10: Sort $\{Item_i\}$ according to $\{Count_i\}$;
11: **return** $\{Item_i\}$;

C. User-based Collaborative Filtering with Social Information

In user-based collaborative filtering, similarity between users is identified by users' purchased items. In this section, a novel method which using users' social information rather than using purchasing records as similarity is proposed. For example, it is believed that two friends have a higher similarity [10], so they will probably have common items. Base on this, for instance, choose a user *A*, and then collect all the items
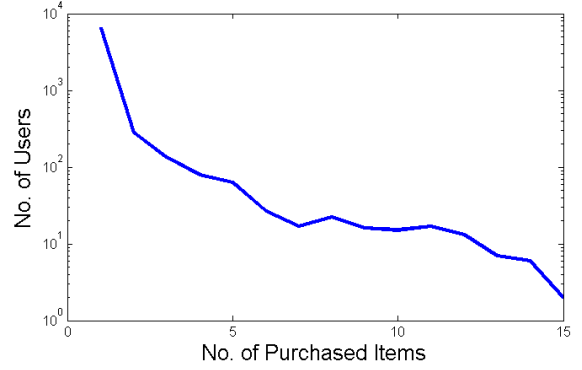
of all the friends of the user *A* (not include items of the user *A*), form an item library. Then count the number of repeated purchases of each item, in other words, how many friends have bought that item. After that, sort the items according to the repeated times and choose top-*N* items to recommend.

IV. EXPERIMENT

A. Dataset

The dataset used in this project is provided by nxTomo, it consist of the information of users, including friendship, levels, leagues, locations, etc. Besides, it also includes the item purchasing history and a record of items in each users pocket.

The data is collected from 7 August 2013 to 16 December 2013, for a more receivable evaluation result, the dataset of purchasing record is split into two parts according to the purchasing date. The earlier one is used for analysis the relationship between items, which is used in making recommendations. The later one is used for evaluate the results of recommendations. The two parts contains almost equal number of records.

Statistically, the dataset contains the information of 197,604 users, and 52,587 purchasing records, which include 7,323 purchased users. The real distribution of how many items each user has is shown in Fig. 4. Most users purchased items amount is under 5, approximately 6,000 users. On the contrast, users rarely have over 15 items.

B. Social Information

In user-based recommendation with social information, first similar users should be find out based on the users profiles. In this paper, 4 types of social relationships has been discussed, respectively are friendships, locations, leagues and levels.

1) Friendship: The given dataset contains 1,615,586 pairs of friends, which means each user has approximately 8 friends on average. It has been proved that friendship indicates a certain similarity, and friends tend to have many things in common.

2) Location: Location is the place that the users registered in the game, which includes 248 places in Hong Kong. Since the regional functions differ from each place, it can been infer that users have same location tend to have similar living and
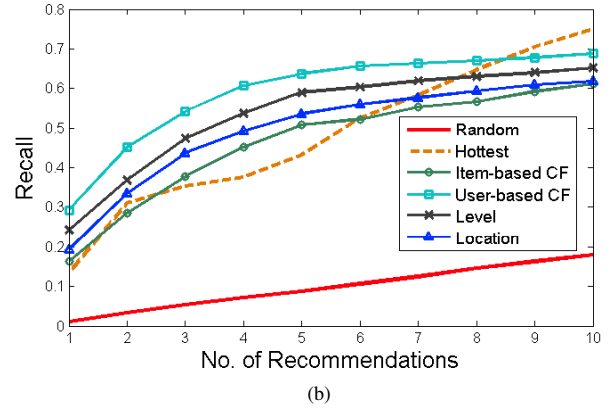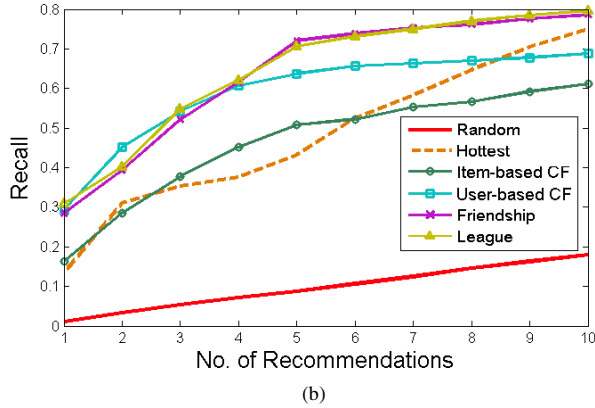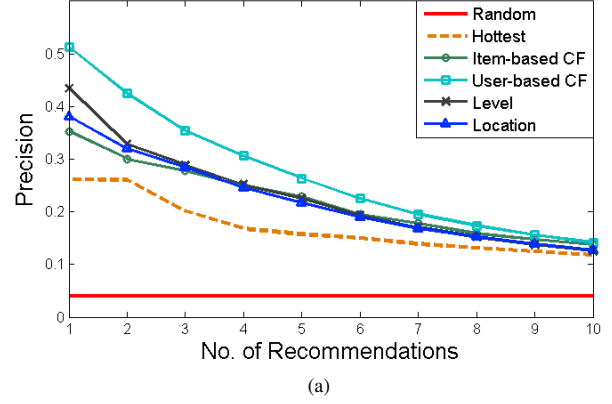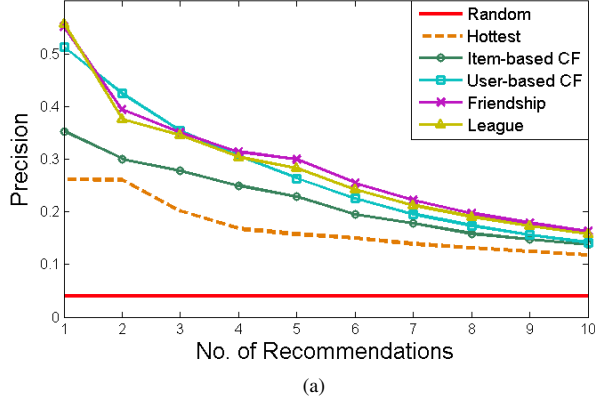
Fig. 5. Evaluation results with friendship and league (a) Precision (b) Recall.

Fig. 6. Evaluation results with level and location (a) Precision (b) Recall.

working environment, so location is also chosen as a type of relationship.

3) League: As a sport game, there are several leagues in the game. Users have to attend league matches in order to join higher leagues. In Barcode Footballer, besides game point, users can also purchase some items by "club fund", which can be only gained from leagues.

4) Level: Similar with league, users in same level will have similarity. Meanwhile, sometimes games will make limitation in games so that only users with lower levels cannot purchase higher level items.

C. Results

To verify the recommendations, a random recommendation and hottest recommendation are made as references. The hottest recommendation curve is drawn by sorting the items according to how many times each item is purchased. Making recommendations by using the above 4 types of relationship respectively, the evaluation result of precision and recall curves of collaborative filtering are shown in Fig. 5 and Fig. 6.

In k-NN based collaborative filtering, it has been proved that higher k values will have higher precision, but this can also cause higher complexity of calculation. In this experiment, the k value is 55 for both item-based and user-based approach. As

Fig. 5 and Fig.6 shows, all the curves of collaborative filtering have higher precision than random and hottest curves, where user-based approach has better performance than item-based approach.

In Fig.5, user-based collaborative filtering using friendship and league have slightly better performance than original user-based collaborative filtering. In Fig.6, the performances of user-based approach with level and location are laid between item-based approach and original user-based approach.

In item-based and original user-based collaborative filtering, the similarities between items or users are calculated from purchasing records, which can cost a long processing time when the data size is large. On the contrast, user-based approach with social information calculate the similarities using users' profiles only, this approach is more efficient than the original method. On the other hand, user-based approaches with both friendship and league as social relationship have better performances than original user-based approach. However, league information needs to be filtered from a large data of users' profiles, while friendship is a separate data table. In terms of efficiency, it is better to use friendship as similar relationship.

V. HYBRID METHOD

In the experiment, it is found that the standard deviation of social information user-based approach is higher than original
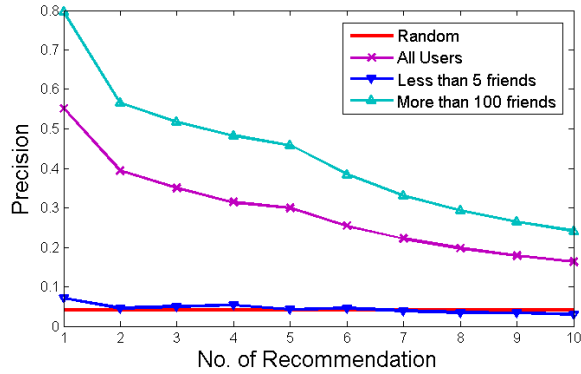
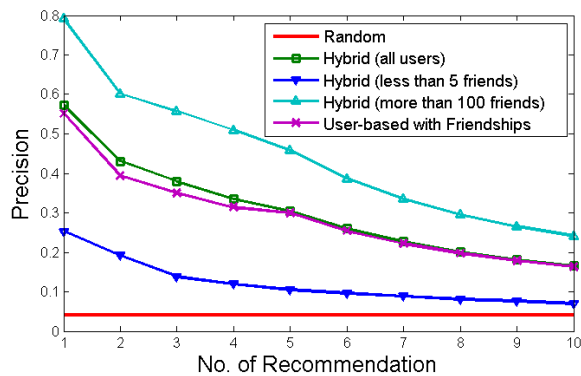Fig. 7. User-based approach with friendship as similar relationship.

Fig. 8. Hybrid approach.

approach, which means the performances between different users have higher variances. Since frinedship have the best performance among all 4 social relationships in social information user-based approach, performance of using friendship is analyzed, as it is shown in Fig. 7, the precision is much lower for those users who have less than 5 friends, while the precision is much higher than average when users have more than 100 friends.

To solve this problem, a hybrid method by combining original user-based approach and social information user-based approach is proposed. As it is described in Section III, a similar user list can be obtained from original user-based approach, and a friends list can be derived from social information user-based approach. Firstly, sort the users' friends list according to the similar users list. Secondly, combine the friends list and similar users list, and then recommend the top-N items from the top-k users. The precision of hybrid approach is shown in Fig. 8, where k value is 55.

As Fig. 8 shows, users with lower friendships using the hybrid method can have a much higher precision than the friendship-based approach. Besides, the overall precision of all users is also higher. For the users who have over 100 friends, since the k value smaller than 100, the recommendation is actually made from users' friends list, the precision does not drop.

VI. CONCLUSION

In this paper, three types of collaborative filtering recommendation methods have been discussed and evaluated in mobile social game items recommendation. As a result, both the user-based method and item-based method has a better performance than random recommendation. However, by making comparisons between these methods, user-based recommendation with friendship as relationship is more accurate and efficient than the item-based and original user-based recommendation. To cope with the decreasing precision when users have a smaller number of friends, a combine method is proposed, which increases the precision at lower number of friends without decreases precision at higher number of friends.

For practical application, in item-based approach, since the gameplay and user habit will not change dramatically, it can be inferred that after an analysis of a certain period of data, the connections between items will keep stable and will not change after adding new data. In user-based approach, since the purchasing records and profiles (i.e. Levels, friendships, etc.) of users will change frequently, it is less efficient to make user-based approach in this scenario, because the system needs to be updated frequently. In hybrid method, both purchasing history and friendship are needed for analysis, it can have better precision and stability but have lower efficiency. In this case, game companies need to trade off between the needed precision and system capacity.

REFERENCES

[1] Adomavicius, G., Tuzhilin, A., "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions." *IEEE Transactions on Knowledge and Data Engineering*, pp. 374-749, 2005.

[2] Viennot, N., Garcia, E., Nieh J., "A Measurement Study of Google Play." *The 2014 ACM international conference on Measurement and modeling of computer systems.*, 2014.

[3] Lettner, F., Clemens, H., "Automated and unsupervised user interaction logging as basis for usability evaluation of mobile applications." *Proceedings of the 10th International Conference on Advances in Mobile Computing & Multimedia.*, 2012.

[4] Yamakami, T., "A two-layer ecosystem evolution model: lessons from stages of mobile data services." *Proceedings of the International on Management of Emergent Digital EcoSystems.* , 2010.

[5] Wohn, D. Y., Na, E., "Virtual item purchasing patterns in a social game: differences between high and low spenders." *Proceedings of the 2012 iConference.*, 2012.

[6] Wohn, D. Y., "Spending real money: purchasing patterns of virtual goods in an online social game." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.*, 2014.

[7] Dutting, P., Henzinger, M., Weber, I., "Maximizing revenue from strategic recommendations under decaying trust." *Proceedings of the 21st ACM International Conference on Information and Knowledge Management.*, 2012.

[8] Leroux, B. Dhoedt, B., Demeester, P., "Performance characterization of game recommendation algorithms on online social network sites." *Journal of Computer Science and Techonology*, 2012.

[9] Herlocker, J., "Evaluating Collaborative Filtering Recommender Systems." *ACM Transactions on Information Systems*, pp 5-53, 2004.

[10]   Drucker, Susan, J., Gumpert, Gary; Cohen, M. Howard, "Social Media." *Regulating Convergence. Communication Law*, 2010.