An Optimization of Collaborative Filtering Personalized Recommendation Algorithm Based on Time Context Information

Xian Jin[1], Qin Zheng[1,2], and Lily Sun[3]

[1]School of Information Management and Engineering, Shanghai University of Finance and Economics, Shanghai, China
[2]South University of Science and Technology of China, Shenzhen, China
[3]School of Systems Engineering, University of Reading, Whiteknights, Reading, UK
13491299@qq.com, qing.zheng@mail.shufe.edu.cn, lily.sun@reading.ac.uk

Abstract. This paper proposes an improved collaborative filtering algorithm based on time context information. Introducing the time information into the traditional collaborative filtering algorithm, the essay studies the changes of user preference in the time dimension. In this paper the time information includes three aspects: the time context information; the interest decays with the time; items similarity factor. This paper first uses Pearson correlation coefficient calculates time context similarity, pre-filtering the time-context. Through the experiment, the improved algorithm has higher accuracy than the traditional filter algorithms without time factor in the TOP-N recommendation list. It proves that time-context information of user's can affect the user's preference.

Keywords: Personalized recommendation, collaborative filtering, time-context.

1 Introduction

Recommender algorithms belong to a class of personalized information filtering technologies that aim to identify which items in a catalog might be of interest to a particular user. The scenario is a very important factor for mobile e-commerce personalized recommendation process. User's interests and needs tend to change with time and space associated with situation changes. Therefore, how to accurately predict the information of interest to users has become a key issue in personalized recommendation problems. Traditional collaborative filtering algorithm does not consider the situation. Therefore, it cannot reflect the needs tendency of users due to the different situations. This shortage caused a lack of personality in recommendation process. The length of time will affect a user memory of interest degree, thereby affecting their user preferences. Meanwhile, the mobile commerce environment is dynamical, and the context information are facing constantly changing, thus affecting the user's interest and purchasing decisions.

Based on this, this paper research on the impaction from time contextual information to users interested and the user interest preferences attenuation occurs over time context

K. Liu et al. (Eds.): ICISO 2015, IFIP AICT 449, pp. 146–155, 2015.

information. This paper provides an UIT (users-project-time) model. The model can achieve precise recommendation and marketing mobile environments. Our research work contributes to three levels: 1) at the method level, we improve the traditional collaborative filtering algorithms, combined with the time context information and collaborative filtering algorithm and improved mobile environment recommended accuracy in traditional collaborative filtering algorithms; 2) at the theoretical level, this study intends to enrich personalized recommendation research of mobile commerce. Currently, most of the research is focused on mobile business mobile business process, trust, risk, and decision-making issue, but consumer behavior, personalized research is still at an exploratory stage. This study also improves the quality of the recommendation system by time context information; and 3) at the commercial value level, the development of mobile Internet and mobile commerce will enter a rapid growth phase, how to provide a valuable service for these users, and then find the right business model is an important strategy for many Internet companies. We provide also a reference model for personalized recommendation in mobile environment.

2 Theoretical Background

2.1 Personalized Recommendation Algorithms

Collaborative filtering recommendation is the most successful technology. For example, Goldberg (1992) describes the similarity between users using artificial mail processing system. Ringo (1997) uses the social information filtering method and the music recommendation system. Currently collaborative filtering is mainly divided into two categories: User-based Collaborative Filtering, and Item-based Collaborative Filtering (Yang, 2012). The User-based Collaborative Filtering algorithm is the oldest algorithms in recommender systems. Goldberg (1992) has applied User-based Collaborative Filtering algorithm into email filtering system, and Grouplens (1994) has then applied the User-based Collaborative Filtering algorithm into news filter. But the key shortage of User-based Collaborative Filtering algorithm is data sparseness; the diversity of recommend list and the scalability of recommend system. To resolve the problems, professor Sarwar (2001) puts forward Item-based Collaborative Filtering algorithm. Amazon, YouTube, and Hulu are using Item-based Collaborative Filtering algorithm.

2.2 Research on time Context Information Recommendation

The traditional personalized recommendation only works with User-Item structure, it does not considerate the context information of user's. It cannot fully reflect the preferences of the user. For example, users like to browse news and information in "morning", like to browse social networks to share information in the “noon”, like to play various entertainment games in the “night”. Some users may order some things without payment online because of the exquisite pictures, but after whole day working, the users feels meaningless and then cancel the order. These user’s preferences changes is due to the context information of user behavior changes. Schilit (1995) divides the context information into three categories, such as computing context, user context, and physical context. Chen (2000) considers that the time context information is another factor. Jonna (2005) divides the context information into five categories, i.e., physical environment, user goals, equipment application, link to person/service, and equipment connection. Kenta (2006) takes views of user's plan, time, user's friends, and external environment in a restaurant recommendation system. Zhang (2007) devises the context information in terms of user preference and context position, such context information refers to the ability to handle user terminal and wireless network. In a music recommendation system, Kim (2008) sets the context information as gender, age, location, time, weather, temperature and pulse for a music recommendation system. Adomavicius and Tuzhilin (2008) put forward the definition of context information and the different method of the user model when context information is constructed. Other scholars (Wang et al., 2012) also offer some research in the theory, methods, and applications of context aware systems, e.g., the New York University, University of Konstanz, the IBM Institute, and the Microsoft research institute. In summary, the diverse context information is appropriately set according to the specific object of research. We select the context information based on the time dimension, and research the change of time context information under user's interest changes, and then improve the traditional collaborative filtering algorithm. With the development of mobile Internet, the research of context information, especially it has great theoretical and practical significance of the context information in mobile environment.

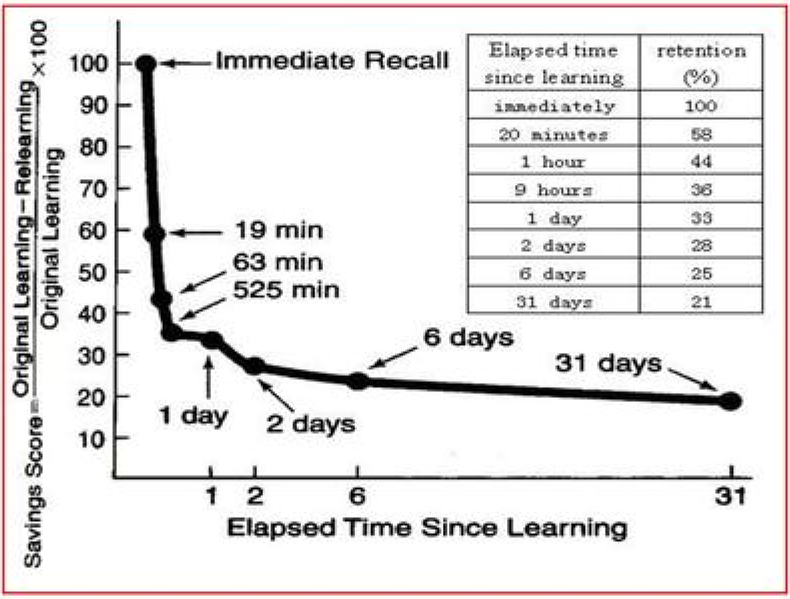

Fig. 1. Ebbinghaus forgetting curve

2.3 User Preferences Attenuation Mechanism Base on Time

Time window mechanism from Song et al. (2006) believes that the user is only interested in recent accessing behavior. Kubat (2006) and Widmer (2007) have improved the time window method that the window size can be automatically adjusted with the change of system prediction accuracy changes. The forgetting function principle is the attenuation of the user's interest and preference by forgetting function. Michalski (2009) and Maloof (2010) use this method to achieve attenuation.

Ebbinghaus (1885) describes the speed of forgetting curve in Figure 1. This curve shows the time weight function is similar to the shape of the forgetting curve forgetting, it is a nonlinear function. The time weight function commonly used for linear function and a nonlinear function (Zheng, 2007; Yin, 2012). Hong and Li (2010) find that recommendation results of nonlinear forgetting functions are better than the linear functions. This paper will use the nonlinear forgotten time functions to improve collaborative filtering algorithm.

3 Methodology

3.1 Data Source and Data Collection

This research work adopts the movies rating dataset of Movielens provided by Minnesota University for our experiment on personalized recommendation. We selected the dataset of Movielens-100K, which contains 100,000 real rating scores for 1682 movies by 943 users. The data sparsity is 6.3%.

This study adopts the accuracy of recommendation as the evaluation index in the other word, the ration of intersection of the TOP-N recommended list and predicting set relative to TOP-N recommended list is better when bigger. The item of predicting set is visited by the user in real world, as defined in Equation (1).

$$p = |\mathrm{List} \cap \mathrm{Test}| \mathrm{N} \tag{1}$$

where N is the count of TOP-N recommended list. N=30.

3.2 Improvement on Time-Base Collaborative Filtering Algorithms

Time context information is introduced to improve the collaborative filtering algorithm on the time dimension. Our study takes the time weight function of users' interest attenuation and the time context weight of the user into consideration in the time to describe the impact on the users' preference, base on which the algorithm used to calculate the similarity between items in the collection can be improve. The following are details about how to decide the weight to improve the algorithm.

3.3 Improvement on time-based weight

Users' preferences decay over time. Compared with the early behavior, the recent behavior of user reflects a better user's interest, based on which, this paper introduce the user's-visit- time-base weight function describing the preference weight changes over time to improve the importance of recent behavior in recommendation. There are two types of weight function: Linear and Nonlinear. Linear function considers the user's preferences are forgetting linearly. Meanwhile the Nonlinear function thinks

changes of user's preferences are nonlinearly. The research by YuHong (2010) shows that the algorithm of closely forgetting curve is superior to the linear function. That's why this paper chooses the nonlinear time function.
Defined by the curve, the attenuation function is a non-increasing function. Time function is defined in Equation (2).

$$f(u,t) = \frac{1}{1+\alpha|tui-tuj|} \tag{2}$$

where
α: decay parameter
tui /tuj : behavior time of user u access the item i/j
f (T) will be smaller when |tui−tuj| is bigger. The value of α can be decided according to the change of system user interest. Lager value means big change since smaller value means small change. The interest degree of user u on item j in the original item based collaborative filtering algorithm is expressed in Equation (3).

$$f(u,\ j) = \sum_{i \in I_u} r_{ui}\, sim(i,\ j) \tag{3}$$

where
Iu means the set of items accessed by user u
j means the predicting item
rui means the interest of user u on item i. rui will be 1 when the user u access the item i. After introducing the time decay function, the interest degree of user u on item i will change over time. Kj means the neighbor items collection of item j. It can be calculated by Equation (4).

$$f(u,\ j,\ t) = \sum_{i \in I_u \cap K_j} sim(i,\ j) \frac{1}{1+\alpha|t_0-t_{ui}|} \tag{4}$$

where
t0 means the current time since tui means the time when user u accesses item i.
α can be decided according to the situation.

3.4 Improvement on Resource-Based Weight

Because the users' early interest is similar with recent interest, this paper takes the similarity between items to measure the similarity between early and recent behavior. T means the recent time period of the user u. IuT is the set of the items accessed by user u in T. Iu means the whole set of the items accessed by user u. We got IuT⊆Iu. P(u,i) means the item-similarity-based weight since i∈Iu. The average similarity between item i∈Iu and the item j∈IuT can be calculated by Equation (5).

$$p_t(u,\ i) = \frac{\sum_{j \in I_{uT}} sim\ (i,\ j)}{|I_{uT}|} \tag{5}$$

where
| IuT | means the count of the items recent accessed in T.

sim(i,j) means the similarity between item i and j, which can be calculated by Pearson correlation coefficient calculating formula:

$$sim(i, j) = \frac{\sum_{u \in I_{ij}}\left(R_{u,i} - \bar{R}_i\right)(R_{u,j} - \bar{R}_j)}{\sqrt{\sum_{u \in I_{ij}}\left(R_{u,i} - \bar{R}_i\right)^2}\sqrt{\sum_{u \in ij}\left(R_{u,j} - \bar{R}_j\right)^2}}$$

By measuring the similarity of the recent and early items accessed, we can figure out the influence on recent interest from the early interest. Sometimes, the interest is forgotten over time. But it can be evoked by the visual or other relevant stimuli, which will impact the users' recent behavior. The resource-based weight can improve the recommendation quality effectively.

3.5 Improvement on Time-Context-Based Information

The mobility of mobile device makes the contextual environment of the user changing dynamically. The interest and requirement will be different in the different contexture environment. For example, user will browser the news app after 18:00 on the metric and shop on line or play games at noon in the office and shop or have dinner outside at the weekend. So that introducing the time contextual information of the user into the recommendation system can help to figure out the users' interest and the weight of each interest.

We take the users' time contextual information into consideration to calculate the similarity of the time to decide the users' interest, which provide the useful information for the recommended results. The vector model was used to express the set of time contextual information collected by n time contextual information. Given TContext= {TC1, TC2, …, TCn},in which TCi={tc1, tc2, …, tcn}, tci($1 \le i \le n$) is a time contextual information at time I and tci is an attribute of the information. This paper category the accesses time by week, and calculate the access preference everyday within a week. The time context information can be described as TC={Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday}. Post to the time context information being defined, User-Time Matrix can be added into time division.

In the User-Time matrix, the task T as item I, the value of T is Ti, (i=0..6). The value in the matrix represents that user's interest degree in a day within a week. Base on this, we can get the time series similarity according to Pearson Correlation Coefficient.

3.6 Calculate User's Interest Preference Base on Items

This research work considers that user's interest preference changes based on time weight, item similar weight and time weight. Then the results can determine user preferences on items, judge the transfer direction of user interest more accurately and then provides a recommendation list more accurate for the users, improve the quality of recommendation system. Firstly the time context information was pre-filtered. Secondly get the time context information similarity set N(t) was retrieved. Finally calculate time weight function and item similarity weight according to the N (t). In the time weight function Tuj$\in$N (t), items of similar time T$\in$ N (t), then u has access to items #i interest preference value user time context information as defined in Equation (6).

$$H(u,t,i) = \alpha 1 \times f(u,t) + \alpha 2 \times pt(u,i) \tag{6}$$

where
αi(i=1, 2)∈[0,1] shows the weight under different factors;
$\sum_i \alpha \mathrm{i} = 1$ f(u,t) means that user #u interested in item #i attenuation degree with time and the time preference;
p(u,i) is the weight based on items similarity;
H(u,t,i)is user interest preference base on the time context information.

The three functions characterize the time context information influence on the user's interests in different aspects, by adjusting the weight values of αi to achieve more accurate recommendations.

4 Results

4.1 The Results of Experiment

Through many times experiments, we obtain the best weight value combination recommendation accuracy rate is 0.2667, then transform the value of TOP-N, validate the recommendation accuracy under different N value. By adjusting the number of items in the list of recommended, determined recommendation accuracy under different N value, the results as shown in Figure 2.

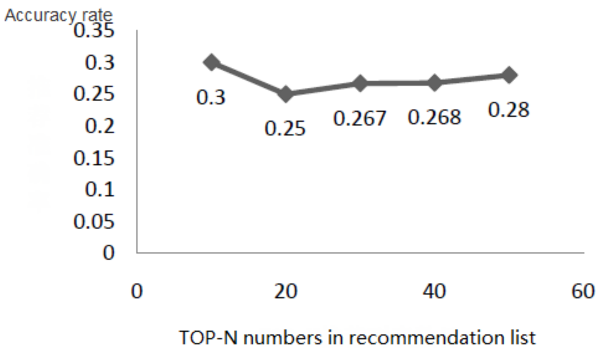

Fig. 2. Recommendation accuracy under Time context information

When N=10, the highest recommendation accuracy rate is 0.3;
When N=20, the highest recommendation accuracy rate is 0.25, achieved at a1>=0.8;
When N=40, the highest recommendation accuracy rate is 0.268, achieved at a1>=0.5;
When N=50, the highest recommendation accuracy rate is 0.28, achieved at a1=1.

Compared with the results produced from Yuhong (2010) and Xing et al. (2007) that do not contain the time context information and other scholars' research, our experimental results show that the recommendation accuracy was significantly higher than the accuracy without time context the comparison of the results, as shown in Figure 3.

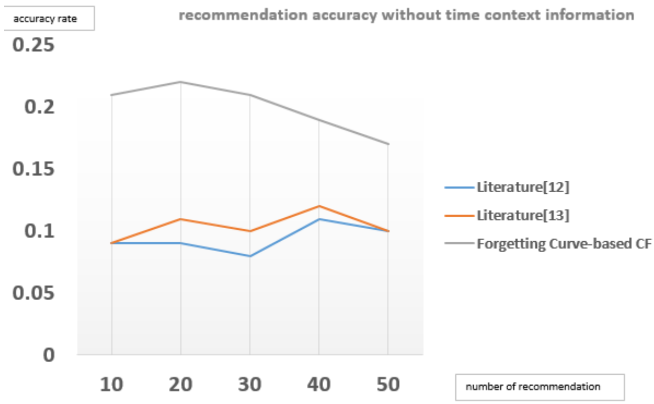

Fig. 3. Recommendation accuracy without time context information

In our result, the recommendation accuracy is no less than 0.25 by improved collaborative filtering recommendation algorithm, even if N value changed in TOP-N recommended list, when N=10, the accuracy rate is the highest for the 0.3. This result is same as most of present e-commerce website recommendation number. Thus, it is obviously shows that users have different preferences under different time context information.

5 Conclusion

Mobile commerce will change people's life. Accessing to the information users' needs in a short period of time become the key value. In this paper mainly results as follows: 1) Apply the time context information to the traditional collaborative filtering algorithms. Established the three dimensional matrix of user-item-time (U-I-T) by pre filtering function and combining the processing time context information; 2) Convert all times information to one day of a week, and then confirm the user's time interest, and then find the item which was accessed the second time at that time point, these results can improve the recommendation accuracy effectively; and 3) Combined the time function and item similarity. The arbitrary time in the interval time functions is the number of days within a week. This method reflects more clearly that user preference is influenced by time context information. On the other hand, combined with the recent information mining. It can determine the impact of the recent behavior of the current user's interests.

The limitation of this paper is that the data source only can shows the visit behavior online. Therefore, the results of this paper needs more validation in actual scenario, such as mobile commerce. That is what I should improve in the future. And, I will continue to research in these problems: (1) accuracy and diversity of the recommendation algorithm. Currently, most of recommendation algorithms only consider recommending accuracy, while ignoring the diversity of the recommendation. But the user's interests are changeable. Therefore, it is lot of research work need to do on how to balance the two fetors the future balance is also a big problem. (2) The different explanation on different data sets. Different datasets shows different characteristics, the performance of various recommendation algorithms on these data sets will also have differences, the reason can be the further studied topics.

Acknowledgements. Our appreciation goes to Ziyi Wang for her valuable comments to this paper.

References

1. Abbar, S., Bouzeghoub, M., Lopez, S.: Context-Aware recommender systems: A service-oriented approach. In: Proceedings of the VLDB Workshop on PersDBLyon (2009)
2. Agrawal, R., Rantzau, R., Terzi, E.: Context-sensitive ranking. In: Proceedings of the 2006 ACM SIGMOD International Conference on Management of Data, pp. 383–394 (2006)
3. Billsus, D., Pazzani, M.J.: Learning collaborative information filters. In: Proceedings of the 15th National Conference on Artificial Intelligence, vol. 4(20), pp. 235–259. AAAI Press, Menlo Park (1998)
4. Breese, J.S., Heckerman, D., Kadie, C.: Empirical analysis of predictive algorithms for collaborative filtering. In: Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence, vol. 461, pp. 143–162 (2008)
5. Chien, Y.H., George, E.I.: A Bayesian model for collaborative filtering. In: Proceedings of the 7th Int. Workshop on Artificial Intelligence and Statistics, vol. 6(1), pp. 152–159 (1999)
6. Dey, A.K.: Understanding and using context. Personal and Ubiquitous Computing 8(10), 256–268 (2001)
7. Getoor, L., Sahami, M.: Using probabilistic relational models for collaborative fi1tering. In: Proceedings of the Workshop Web Usage Analysis and User Profiling, vol. 9(12), pp. 134–147 (1999)
8. Herlocker, J.L., Konstan, J.: Content-independent task-focused recommendation. IEEE Internet Computing, 240–257 (2001)
9. Kim, J.H., Jung, K.Y., Ryu, J.K.: Design of Ubiquitous Music Recommendation System. In: Proceedings of the 4th International Conference on Networked Computing and Advanced Information Management, pp. 250–263 (2008)
10. Jani, J.: Collaboration in Context-Aware Mobile Phone Applications. Proceedings of the 38th Hawaii International Conference on System Science 5(13), 156–172 (2005)
11. Oku, K., Nakajima, S., Miyazak, J.: Context-aware SVM for context-dependent information recommendation. In: Proceedings of the 7th International Conference on Mobile Data Management, vol. 5(19), pp. 189–207 (2006)
12. Kobsa, A.: Generic user modeling system: user modeling and user-adapted interaction. Information Research System 9(12), 143–165 (2001)
13. Melville, P., Mooney, R.J., Nagarajan, R.: Content-boosted collaborative filtering for improved recommendations. In: Proceedings of the l8th Conference on Artificial Intelligence, vol. 5(15), pp. 243–264 (2002)
14. Park, H.S., Yoo, J.O., Cho, S.B.: A context-aware music recommendation system using fuzzy Bayesian networks with utility theory. In: Proceedings of IEEE Fuzzy Systems and Knowledge Discovery (2006)

15. Pavlov, D., Pennock, D.: A maximum entropy approach to collaborative filtering in dynamic, sparse and high dimensional domains. In: Proceedings of the 16th Annual Conference on Neural Information Processing Systems (2002)
16. Pazzani, M., Billsus, D.: Learning and revising user profiles: the identification of interesting web sites. Machine Learning (1997)
17. Stender, M., Ritz, T.: Modeling of B2B mobile commerce processes. International Journal of Production Economics 9(19), 359–372 (2006)
18. Hong, Y., Yun, L.Z.: Collaborative filtering recommendation algorithm based on forgetting curve. Journal of Nanjing University 46(5), 34–38 (2010)
19. Chunxiao, X., Fengrong, G., Zhan, S.N.: A collaborative filtering recommendation algorithm to adapt to the user interest change. Journal of Computer Research and Development 44(2), 10–14 (2007)