Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2015, 7(3):2487-2490

Research Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

Study on personal preferences mining method in O2O E-commerce model

Shang Peini and Yuan Feiyun

College of Information Engineering, Yu Lin University, Yu Lin City, Shaanxi, China

---

ABSTRACT

In this paper, we studied the theory of personalized recommendation, compared the different recommendation technologies, and analyzed the applicability of several recommendation technologies to the O2O mode of e-commerce. And then an approach for mining personal context-aware preferences from the context-rich device logs, or context logs for short, and exploit these identified preferences for building personalized context-aware recommender systems was proposed. The experiment results show that the recommendation method put forward in the thesis is able to meet the demand of the O2O e-commerce mode.

Key words: O2O E-commerce, Personal preferences, Data Mining, user data

---

INTRODUCTION

With the rapid development of e-commerce, O2O e-commerce, as a new and efficient business model, rapidly spreads. Along with the appearing of online group-buying websites, more and more companies focus on the O2O mode of e-commerce. Nevertheless, as the situation of data exploded, a huge amount of data makes the data overload and redundant, which usually makes the users get lost in the sea of data and be difficult in finding the product successfully they need. Against this background, personalization technology, as a key technology of e-commerce, is becoming one of the research focuses in the area of O2O electronic commerce. However, there are relatively few in-depth studies on personalization technology in the area of O2O electronic commerce, and the accuracy and efficiency of these recommendation algorithms still needs to be improved.

In recent years, although many researchers have been focusing research on personalized context aware recommendation problem and proposed some mining mobile users personalized context preference method [1-3, 4, 5]. However, these methods only consider the context information usually single, at the same time, most of the work platform of Internet based project scoring record. In fact, the mobile user context log user activity recorded with rich semantic information can help more accurately mining personalized user preferences in different complicated situations .And how to through the personalized preferences to learn mobile user context log mining and construction of personalized context aware recommender system is still an urgent need to solve the open problem.

Recently, a lot of mobile user oriented personalized context aware recommender systems research is published in the academic circles. Tung et al [6] in tourism information service proposed a personalized mobile consider user context information recommendation system. Park et al. [5] proposed a recommendation system based on the geographical location of the personalized, by the Bayesian network (Bayesian Networks) modeling context information to mobile users, which found that the personalized user preference. [7] Bader et al proposed a new context aware method to recommend a point of interest for mobile users (Points-of-Interest). Woemdl et al. [8] proposed a "play.tools" hybrid framework to implement mobile user context information recommendation based on App. This recommendation framework assumes a mobile App mobile users will like other users to install in the current context. Kim et al. [4] Research of many collaborative filtering based on (Collaborative Filtering CF) method to construct based on context aware ad recommendation. Unlike the methods mentioned above, the framework of the method proposed in this paper can be modeled explicitly in the context information and user activity categories, implicit conversion without scoring record.

EXPERMENTAL SECTION

The experimental data set

In this paper, experiments are carried out using MovieLens data. MovieLens is one of the most famous data sets the field of recommender systems, the GroupLens mechanism is responsible for the establishment of the entire MovieLens web site and data collection. It can be downloaded by the address http://grouplens.org/datasets/movielens/ , GroupLens to the academia provides 100k, 1M and 10M three the size of the user item rating data set, with a memo information data set. MovieLens100K- consists of 1000 users on 1700 film score of 100000. MovieLens1M- includes 4000 to 1000000 film scores of the 6000 users. Application of MovieLens10M- including 72000 users to 1 thousands of evaluation of 10000 films, 100000 tag data.

Evaluation standard

Assessment criteria recommended performance measure includes two kinds: online and offline evaluation. In the aspect of recommendation performance measure personalized recommendation algorithm, this paper chooses the prediction accuracy rate.

The accuracy of prediction evaluation criteria include statistical accuracy (Statistical Accuracy Metrics) and decision support accuracy (Decision Support Accuracy Metrics). Statistical precision refers to the error of the result and the real evaluation of recommendation. Decision evaluation can also be called classification accuracy measurement method, refers to the right to determine whether the user likes some proportion of project.

Statistical precision mainly include MAE (mean absolute error), RMSE (standard error), NMAE (mean absolute error criterion) etc..

The experimental environment

The operation environment of the program are as follows:
Memory: 2 GB
Hard disk: 300 GB
Operating system: Microsoft Windows XP
CPU:Pentium (R) Dual-Core CPU E5300 @ 2.6GHz
The use of the software: Microsoft Office Excel 2010, Microsoft Visual 2010, MATLAB Studio 2010a and Eclipse.

RESULTS AND ISCUSSION

Experiment 1

The use of MovieLens data sets, different number and different nearest neighbor clustering center number under the condition of test performance, improved collaborative filtering algorithm based on fuzzy clustering. Neighbor count recently from the beginning of the 5, growth step is 5, the maximum is 40. The clustering center number from 5, increased to 20, the interval is 5. Under other factors under the same conditions, the calculation of the MAE value. The results are shows in figure 1.

As shown in figure 1. Through a number of control changes the cluster center, the obtained results may vary greatly, not only affects the accuracy of recommendation, changes will affect the whole curve. Select a different number of nearest neighbors is different, also can get the accuracy of recommendation different, the degree of change curve is also affected by the. Integrated two influencing factors, can be drawn from the chart, the clustering center is 20 cases, the accuracy of recommendation representatives recommended value is the highest, the worst results. When the clustering center number is 5, the accuracy of recommendation value is small, the recommendation result more ideal, is also the recommendation accuracy and the clustering center is inversely proportional to the number of. Through the curve, with the nearest neighbor a recommended value is relatively small in between 10-20, the curve changes slowly, from the nearest neighbor number 20 began, the accuracy of recommendation value is greater, the worse the recommendation effect. Therefore, the number of nearest neighbors should be less than 20, the accuracy of recommendation will be relatively high.

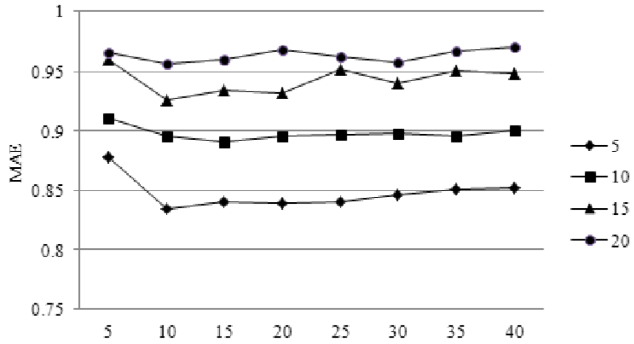

Fig1. **Experimental results figure**

Experiment 2

Different sparseness of data sets with different number of nearest neighbor case, the mean absolute error of the improved algorithm test. In the other test under the same conditions, the selection of MovieLens data set according to certain standards are divided into different subsets of the data. General in machine learning, will use the k- fold cross validation is used to divide the data, training out most of the data, a small part of the remaining data is used to calculate the error, can be used K times calculated as the average of the evaluation criteria. This paper chooses 4-fold cross validation experiments, 4 groups of test after the experimental results as shown in figure 2.

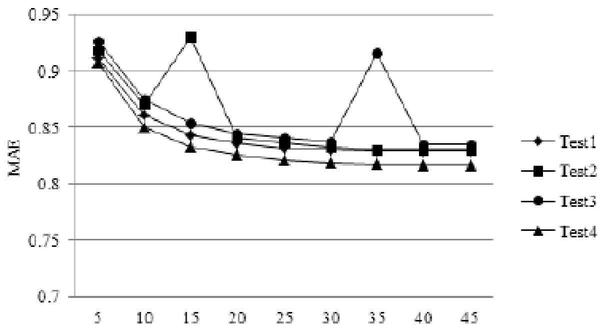

Fig2. **Different sparsity comparison chart**

As can be seen from the figure, a number of nearest neighbor influence each curve trend, different also greatly influence the value of the MAE. Compared traditional collaborative filtering algorithm improved algorithm comparison and presented in this paper based on clustering, the nearest neighbor number less than 20, the improved algorithm value smaller than traditional algorithm in MAE, which recommended the accuracy. Although after the nearest neighbor is greater than 20, the improved algorithm MAE value will be greater than the traditional algorithm, but the gap is not large, the recommendation system in O2O electronic commerce website, in the search for a nearest neighbor number under the condition of low, set recommended nearest neighbor number is less than 20, in the line recommended shorter time-consuming calculations, real time better clustering collaborative filtering algorithm, the recommendation accuracy is superior to the traditional fuzzy clustering improved recommendation algorithm.

CONCLUSION

This paper systematically analyzed the O2O electronic commerce model, and then studied the characteristics of data in O2O electronic commerce, expounds the relationship between e-commerce and personalized O2O. Then compares the O2O model and B2C model, focuses on the O2O model. Finally put forward the recommendation system structure of O2O electronic commerce mode and its recommendation model structure, and the set of test and results of the recommendation algorithm improved analysis using MovieLens data. The experimental results show that the improved algorithm of recommendation efficiency and the accuracy of recommendation are improved.

Acknowledgments

This research was financially supported by the special research of the Yulin Univesity (Grant NO.15YK29).

REFERENCES

[1] Zheng V W,Cao B,Zheng Y,et al. , *Proceedings of the Twenty-Fourth AAAI conference on Artificial Intelligence,* **2010**,236-241.

[2] Ge Y, Liu Q, Xiong H, et al., *Proceedings of Proceedings of the 17th ACM SIG KDD international conference*

on Knowledge discovery and data mining,New York, NY, USA,ACM, **2011**,983-991.
[3] Liu Q,Ge Y' Li Z, et al., *Proceedings of the 2011 IEEE 11th International Conference on Data Mining*, Washington, DC, USA, IEEE Computer Society,**2011**,407-416.
[4] Kim K, Ahn H, Jeong S, *Journal of World Academy of Science, Engineering and Technology*, 2010(64) ,357-362.
[5] Park M H, Hong J H, Cho S B, *Lecture Notes in Computer Science*, Springer, **2007**,1130-1139.
[6] Tung H W, Soo V W, *Proceedings of the 2004 IEEE International Conference on e-Technology, e-Commerce and e-Service (EEE'04)*, Washington, DC, USA,IEEE Computer Society,**2004**, 259-262.
[7] Bader R, Neufeld E,Woemdl W, et al. , *Proceedings of Proceedings of the 2011 Workshop on Context-awareness in Retrieval and Recommendation*,New York, NY, USA: ACM, **2011**,23-30.
[8] Woemdl W, Schueller C, Wojtech R,*Proceedings of Proceedings of the 2007 IEEE 23rd International Conference on Data Engineering Workshop*, Washington, DC, USA: IEEE Computer Society, **2007**,871-878.