International Journal of Grid and Distributed Computing
Vol. 10, No. 1 (2017), pp.255-268
http://dx.doi.org/10.14257/ijgdc.2017.10.1.22

Research on Collaborative Filtering Recommendation Algorithm based on User Interest for Cloud Computing

He Kun

College of computer science and technology, Changchun Normal University, Changchun 130032, china
178381448@qq.com

Abstract

Aiming to resolve the mobile commerce scenario suggested problems, cloud computing technology and mobile user context are combined to propose a collaborative filtering model based on user interest in mobile scenarios. Through computing the scene similarity based on mobile users, we find similar scenarios constructed target user set current situation, and then establish the item scoring scene and scoring matrix. Based on Map Reduce, we propose a collaborative filtering recommendation method to realize parallel recommendation.

Keywords: *mobile commerce, recommendation model, cloud computing, Map Reduce*

1. Introduction

At present, lots of scholars in China and overseas are dedicated to studying the issue of user interest mining and service recommendation in mobile environment, e.g. mobile-based news browsing and information retrieval; music oriented to mobile business users, TV program and news recommendation [1-3]. But many of current mobile business information providers can hardly meet people's needs of them above. The reason can be on the one hand, a majority of mobile business information operators provide to user the mobile package like recommending to tourists a series of combined service including scenic spots, travel guide, traffic, hotels and vehicle rental. However such service package can't satisfy to their heart because the package is tailored as per their actual needs, in other words, it doesn't consider user's individualized requirement; on the other hand, such service package is just a simple assembly of plentiful service information, rather than customized recommendation in accordance to user's personalized preference, user's circumstances and user's economically situation [4-6]. Hence each partition of such package is mutually separated [7]. In that case, user can only take advantage of information filtering technology like search engine to discover from tremendous service data the most relevant items in line with own actual situation [8-9].

As the most widely used individualized retrieval technique, collaborative filtering recommendation predicts user's degree of preference for some item based on user itself and item rating data [10-12]. But, in the course of the mobile business service recommendation, it's far from enough to consider only user and item' scoring data; instead, it needs take into account the circumstances of mobile users, such as serving time, geographic location, season/weather/temperature of the time, user's social role and historic preference [13-15]. As in the process of tourist attractions Recommended, the bad weather for travelers to recommend indoor tourist attractions. Passengers want to find attractions in the vicinity of the area, it is not recommended to the scenic spots. That is to fully consider the situation of passengers at the time recommended. In addition, other users should also consider the evaluation of the item.

ISSN: 2005-4262 IJGDC

In the same case, a case study of passenger oriented near scenic spot recommendation. when the user A tours a scenic spot, and then to the attractions of satisfaction evaluation. It implies that it is the time when the tourist attractions, such as the time of play, the season of the time and the climate.

When the recommended system for the user B recommended its interest in tourist attractions, if B and A has a strong similarity at this time, so, the user A at the time of the evaluation of information on the user B has a strong impact.

So, one of the pressing difficulties to be solved in mobile recommendation is to introduce to recommendation system user’s situation interest for individualized modeling as to provide user with the most relevant recommendation service with contexts of the time, in light of mobile business’ sensitivity to location and situation [16-19].

In order to solve the above problems, in the research of Mallat puts forward the concept of knowledge management to manage the user's situation information, and according to the user's personalized scene to provide personalized service.

Adomavicius in the personalized recommendation will be the user's purchase time as a scenario variable. Then combined with the traditional collaborative filtering recommendation method, a multi-dimensional recommendation model is constructed. In order to make up for the deficiency of the traditional collaborative filtering method, Li Deyi and Zhang Guangwei considered the impact of item classification on user rating when they were in the prediction of scoring, a recommendation model is based on the construction of the item scenario.

In addition, the combination of the neural network and collaborative filtering recommendation model is used to measure the similarity among the items by using the basic information of the users, construct the score similarity matrix for collaborative recommendation, and improve the quality of the system. Tao Mengying and Zhou Puxiong put forward the multi-level three-dimensional recommendation model of “LBS+AR” multilayer association rules for the mobile service, the accuracy and quality of the system are improved by introducing the scene sensitivity calculation.

Yuan Jing and Jiao Yuying's research also show that the application of mobile commerce scenario information to personalized recommendation system is a major trend in the future research. On the basis of this, a Learning Resource Recommendation Model Based on user scenario is established, and the learning resources are sorted by user's situation information.

To sum up, the recommendation model based on user's situational interest has attracted wide attention of scholars both at home and abroad. Some scholars introduce user scenarios into the mobile commerce recommendation system.

However, the current research is based on the user's static information or classification of the classification of the item to classify users, lack of comprehensive mining of mobile commerce user's location, time and business needs. In addition, the user behavior data in the mobile environment is exponentially increasing, which greatly increased the difficulty of finding useful knowledge and related interest from massive data.

Therefore, aiming at the problem of personalized recommendation of mobile commerce, in this paper, a collaborative filtering recommendation model based on mobile commerce user scenario interest is proposed, which is introduced into the collaborative filtering recommendation algorithm. Then the MapReduce programming model of cloud computing is used to realize the collaborative filtering algorithm. Finally, the recommendation results are given.

2. Collaborative Filtering Recommendation Model of Mobile Business Based on User's Situation Interest

2.1. Traditional Collaborative Filtering Recommendation Methods

The core thinking of collaborative filtering recommendation system is to find out given user's similar (interest) user from user groups by analyzing user interest; then the system will have prediction of the user's degree of fondness of certain information through combining ratings of the information by those similar users. But with rapid development of Internet, traditional collaborative filtering recommendation system faces a skyrocketing number of users. The collaborative filtering recommendation based on user interest mining not only needs to calculate the similarity among numerous users but also is challenged with online calculation of new coming users of a big quantity. Among existing recommendation systems, the commonest recommendation pattern is collaborative filtering recommendation based on item. The item-based collaborative filtering recommendation is capable to compute offline the similarity among rated items, which is helpful to enhance the response capability of the entire system. The specific steps of the method are as follows:

Step1: Establish the degree of recommendation score matrix of the User and Item. Before establishing the scoring matrix, the User is calculated by using the method of pre setting weighted method to calculate the recommendation degree of Item in the view, stop and buy items. It is shown in Table1. Since the user is not likely to experience all of the items, so the matrix part of the score is empty.

Table 1. User-Item Score Matrix

$R_{user,item}$	User1	User2	User3	User4
Item1	$R_{1,1}$	$R_{2,1}$	$R_{3,1}$	$R_{4,1}$
Item2	$R_{1,2}$	$R_{2,2}$	$R_{3,2}$	$R_{4,2}$
Item3	$R_{1,3}$	$R_{2,3}$	$R_{3,3}$	$R_{4,3}$

Figure 1. Similarity Degree of Collaborative Filtering Recommender Model

Step 2: According to the known value of the matrix of the degree of recommendation, according to the specific similarity calculation method, the Pearson coefficient method is used to calculate the similarity between each Item, thus build the most similar items list for each Item. It is shown in Figure1.

Step3: According to the item score and the most similar items list score information, according to the specific rules makes score prediction.

Step4: Summarize all the items in the forecast value and arrange according to the numerical value. Take the maximum value of the first N item (that is, Top-N) as a result of the recommended output.

2.2 Situation Interest Description of Mobile Business User

When collaborative filtering recommendation technique is applied to do recommendation, most recommendation systems take rating matrix which is based on user-item to calculate recommendation result. As mentioned above, it's not the most accurate to do recommendation by only referring to information about user and item; instead, it requires to consider other elements like user's location, season of the time, time of the zone, social information etc. as to comprehensively dig out user's potential interests. In this part, firstly it introduces the situation interest of mobile business user; on that basis, the situational model will be created for mobile business user rating; then it discusses the degree of similarity among user situations.

The mobile business rating situation proposed here refers to actual situational information involved with item rating after mobile business user experiences one specific item, such as geographic location, time situation, social situation where when the user gives points. Here it firstly defines user's item rating situational model. The situational information in the recommendation system has n kinds. In the paper, it will describe formally item rating situational model in the expression of vector.

$$ItemRatingContext = (Item, Context_1, Context_2, ..., Context_n)$$

The target user's situation information uses $ItemContext$ to express, it has been experienced in the item and to make a score of the user rating scene recorded as $ItemRatingContext$,the similarity of $ItemContext$ and $ItemRatingContext$ is calculated:

$$Sim(ItemContext, ItemRatingContext) = \frac{1}{n}\sum_{i=1}^{i=n} Sim_i(ItemContext, ItemRatingContext) \quad (1)$$

Where, $Sim(ItemContext, ItemRatingContext)$ is the similarity between $ItemContext$ and $ItemRatingContext$, to represent the similarity between the target user's scene and the rating of the user.

The similarity between the two in a particular scenario type i is expressed as $Sim(ItemContext, ItemRatingContext)$, if $Sim(ItemContext, ItemRatingContext) > T$,then it shows that the score has a high degree of similarity between the user and the target user. That is to obtain a higher recommended weight.

2.3. Collaborative Filtering Recommendation Model Based on Situation Interest Rating Matrix

The kernel algorithm of the proposed collaborative filtering recommendation model of mobile business based on situation interest is the collaborative filtering approach based on situation interest rating matrix, i.e. combing item rated situation and rating matrix to choose specific mobile business recommended service item. Its main steps are as follows:

1 Acquire information about user-item rating matrix and rated situation

Establish the collaborative filtering model based on situation interest rating matrix; firstly create user-item rating matrix $RS_{M\times N}$; matrix $RS_{M\times N}$ is User's scoring of Item. The formal expression is shown in formula (2).

$$RS_{M\times N}=\begin{pmatrix} rs_{u_1s_1} & rs_{u_1s_2} & \dots & rs_{u_1s_N} \\ rs_{u_2,s_1} & rs_{u_2,s_2} & \dots & rs_{u_2,s_N} \\ \dots & \dots & \dots. & \\ rs_{u_Ms_1} & rs_{u_M,s_2} & \dots. & rs_{u_M,s_N} \end{pmatrix}_{M\times N} \tag{2}$$

2 Calculate similarity between target user and others

When the similarity is calculated between target user and other users, it uses the most widely used Pearson correlation coefficient measuring formula; with Pearson correlation coefficient method, the degree of similarity between target user iu and user ju can be estimated. The formal expression is shown in formula (3).

$$sim(u_i,u_j)=\frac{\sum_{s_k\in US_{u_iu_j}}(rs_{u_i,s_k}-\overline{rs}_{u_i})(rs_{u_j,s_k}-\overline{rs}_{u_j})}{\sqrt{\sum_{s_k\in US_{u_iu_j}}(rs_{u_i,s_k}-\overline{rs}_{u_i})^2}\sqrt{\sum_{s_k\in US_{u_iu_j}}(rs_{u_j,s_k}-\overline{rs}_{u_j})^2}} \tag{3}$$

3 Ultimate rating prediction and Top-N recommendation

As per steps of collaborative filtering recommendation method, the modified similarity measuring formula sim is used to select the N users who share the highest similarity as similar user collection; then do weighting of similarity between target user and similar user to ultimately get prediction rating. The formal expression is shown in formula (4).

$$P_{u_i,s_j}=\overline{rs_{u_i}}+\frac{\sum_{u_p\in SU}(rs_{u_p,s_j},-\overline{rs_{u_p}})\times sim''(u_i,u_j)}{\sum_{u_p\in SU}sim''(u_i,u_j)} \tag{4}$$

3. Filtering Recommendation Method in MapReduce Framework

Collaborative filtering recommendation method itself has certain drawbacks, which cause plenty of recommendation systems to become weak in scalability. Although researches were done to improve that through certain optimized strategy, most are only optimization of the algorithm and run on single server. The mobile business user situation interest recommendation discussed here is oriented with tremendous mobile business user behavior data, of which user situation data cannot be effectively explored merely by single machine improving collaborative filtering approach. On that basis, the paper introduced advanced cloud computing processing technology. First of all, it builds Hadoop distributed cloud environment to deploy recommendation model; then it realizes parallel recommendation with the collaborative filtering recommendation in MapReduce framework so as to solve the issue of expandability and scalability of traditional collaborative filtering recommendation system.

3.1 MapReduce Technology

MapReduce is a kind of parallel and distributed computing model developed by Google, it has been widely used in the field of search and processing of massive data. Map is a decomposition process, which will be distributed to a computer each data fragment processing, so as to have the characteristics of the distributed. Reduce is to separate the data to be integrated together, and finally the results will be aggregated output. The main processes of MapReduce include:

Step 1: when using the MapReduce function, the input file is decomposed into M tasks.

Step 2: the main control program Master is responsible for the scheduling and monitoring of the program, is the ruler of the overall situation. It finds out the idle Worker node assignment Map R tasks and Reduce M tasks need to be assigned to each computing node.

Step 3: enter a processing task into a Map job, it first pre-process the input data, and will be separated from the file input, after processing the key word Key, passed to the Map function.

Step 4: The intermediate result generated by the Map function automatically writes the intermediate result to the local disk in a period of time, and the timing of the storage of the results to be passed to the host program Master, then Master is responsible for the storage of the specific location of the information transmitted to the Reduce sub task node.

Step 5: the node that performs the Reduce subtask gets the task from the Master, call remote procedure from the Map working on the local hard disk to read buffer intermediate data. When the Reduce machine reads all of the intermediate data, it uses the intermediate key to sort it out. This information with the same key will eventually be aggregated together.

3.2 Collaborative Filtering Method Based on MapReduce

During the collaborative filtering recommendation, there are two core steps:

(1) Calculation of similarity based on user-item rating matrix;

(2) Rating of unrated item based on similarity prediction. The two steps are implemented in certain sequential order so that it can be regarded as two serial steps. Especially when the similarity is calculated based on user-item rating matrix, any two rated items' similarity calculation is an uncoupling and parallel process. That is in compliance with the distributed parallel treatment idea in cloud environment and can be implemented through MapReduce tasks. The rating procedure for predicting unrated items can be also implemented by MapReduce. Next it will calculate similarity based on user-item rating matrix together with MapReduce workflow; then, foresee the rating value of unmarked item.

As known from the definition of MapReduce programming model, one MapReduce is actually a processing course involving the acquisition of output key value pair by means of input key value pair. When the similarity of user-item rating matrix is computed, input key value pair is expressed like <null，(User ，Item，Score)>; output key value pair like <(Item1，Item2)，Sim>.

User-item rating matrix similarity calculation is completed by two MapReduce tasks. The first MapReduce is adopted to collect scoring of items by users then rank in order as per user name; Map function at this phase converts input data to relative key value pair; then use Reduce function to merge items which belong to the same user. It is shown in table1-table2.

The second MapReduce is adopted to estimate degree of similarity among items; key value pair between user and each item is transformed to item-item key value pair; that is achieved in this way: Map function obtains comparison of ratings by the same one User between each Item; meanwhile, use Reduce function to reckon similarity among items. It is shown in table3-table4.

Table 1. The Map Phase of the first Map Reduce

Map input	Map output
<null,(User1,Item1,Score)>	< User1,(Item1,Score)>
<null,(User1,Item2,Score)>	< User1,(Item2,Score)>
<null,(User1,Item3,Score)>	< User1,(Item3,Score)>
<null,(User2,Item1,Score)>	< User2,(Item1,Score)>
<null,(User2,Item2,Score)>	< User2,(Item2,Score)>
<null,(User3,Item2,Score)>	< User3,(Item2,Score)>
<null,(User3,Item3,Score)>	< User3,(Item3,Score)>

Table 2. The Reduce Phase of the First Map Reduce

Reduce input	Reduce output
< User1,(Item1,Score)>	< User1,((Item1,Score), (Item2,Score), (Item3,Score))>
< User1,(Item2,Score)>	
< User1,(Item3,Score)>	
< User2,(Item1,Score)>	< User2,)(Item1,Score), (Item2,Score))>
< User2,(Item2,Score)>	
< User3,(Item2,Score)>	< User3,((Item2,Score), (Item3,Score))>
< User3,(Item3,Score)>	

Table 3. The Map Phase of the Second Map Reduce

Map input	Map output
<User1,((Item1,Score),(Item2,Score), (Item3,Score))>	<(Item1,Item2),(Score, Score)> <(Item2,Item3),(Score, Score)>
<User1,((Item1,Score),(Item2,Score), (Item3,Score))>	<(Item1,Item3),(Score, Score)>
<User1,(Item2,Score),(Item2,Score))>	<(Item1,Item2),(Score, Score)>
<User3,(Item2,Score),(Item3,Score))>	<(Item2,Item3),(Score, Score)>

Table 4. The Reduce Phase of the Second Map Reduce

Reduce input	Reduce output
<(Item1,Item2),(Score,Score)> <(Item2,Item2),(Score,Score)>	<Item1,Item2), Sim>
<(Item1,Item3),(Score,Score)> <(Item2,Item3),(Score,Score)>	<Item1,Item3), Sim>
<(Item2,Item3),(Score,Score)>	<Item2,Item3), Sim>

After two MapReduce treatment, the similarity calculation results were obtained, and the similarity of each Item was collected and sorted. It is shown in table5-table6.

Table 5. The Map Phase of the Third Map Reduce

Map input	Map output
<(Item1,Item2), Sim >	<Item1,(Item2,Sim)> <Item2,(Item1,Sim)>
<(Item1,Item3), Sim >	<Item1,(Item3,Sim)> <Item3,(Item1,Sim)>

Table 6. The Reduce Phase of the Third Map Reduce

Reduce input	Reduce output
<(Item1,Item2), Sim >	<Item1,(Item2,Sim),(Item3,Sim))>
<(Item1,Item3), Sim >	
<(Item2,Item1), Sim >	<Item2,(Item1,Sim),(Item3,Sim))>
<(Item2,Item3), Sim >	
<(Item3,Item1), Sim >	<Item3,(Item1,Sim),(Item2,Sim))>
<(Item3,Item2), Sim >	

Finally, according to the calculation of the user's recommendation score similar list, use the Map function to predict. Finally use the Reduce output recommended results, it is shown in table7-table8.

Table 7. The Map Phase of the Fourth Map Reduce

Map input	Map output
<Item1,(Item2,Sim),(Item3,Sim))>	<null,(Item1,PreScore)>
<Item2,(Item1,Sim),(Item3,Sim))>	<null,(Item2,PreScore)>
<Item3,(Item1,Sim),(Item2,Sim))>	<null,(Item3,PreScore)>

Table 8. The Reduce Phase of the Fourth Map Reduce

Reduce input	Reduce output
<null,(Item1,PreScore)>	<null, List of recommended items>
<null,(Item2,PreScore)>	
<null,(Item3,PreScore)>	

4. Experiment Design and Discussion

4.1. Experimental Environment

In simulated cloud computing environment, used 9 servers to build a Hadoop environment, one of which as a NameNode, the rest as DataNode, the specific construction process includes the following steps:

Hadoop cloud environment system software is installed on the VMware virtual machine system. Virtual machine Linux Linux system used 5.5-x64 jdx, jdx1.60 used RedHat version. Hadoop used the -0.21.0 Hadoop-0.21.0 version.

Hadoop cloud computing environment to build 9 server hardware configuration is: Lenovo server, where, the memory is 4G, hard disk is 1T, CPU frequency is 2.8G. The other 8 servers used the common PC, memory is 2G, hard disk is 320G, CPU frequency is 2.6G.

Hadoop cloud environment of NameNode is the Lenovo server, named Hadoop; the remaining 8 units as Hadoop DataNode PC, specifically named Hadoop 1, Hadoop 2,... Hadoop 8.

After the construction of the basic Hadoop architecture, the need for Hadoop configuration and installation, so as to achieve the user no password access of the Hadoop nodes, NameNode and DataNode nodes are without a password access, and test the Hadoop configuration can be successfully run.

4.2Selection of Dataset

In the course of validating recommendation model, whether the testing dataset which is adaptive to the algorithm can be selected is critical to whether an ideal testing effect can be achieved. Up to now fewer papers have been found about situation recommendation; open and available situation recommendation dataset is deficit and thus there are too fewer data sets about situation information oriented to mobile environment. During experimentation, the paper used Moviepilot dataset and MobileServices dataset stated in [20]. The former is oriented to movie recommendation with contexts of situation, containing time situational information when users are scoring movies (rating point is integer in 0-100). So from Moviepilot original dataset, we chose some users and those who did reviews for the most times as testing set for the experiment, which includes 1 628 955 pieces of rating records by 2320 users for 23628 movies.

MobileServices dataset is the one of situation recommendation oriented to mobile environment. The paper [20] captured data from online application store of China Mobile; then it investigated and analyzed mobile business user behavior research report. Based on that, a mobile network service recommendation system and simulated data integration platform was devised. The data integration platform stores multiple mobile business user behavior data and situation information. Here it employs MobileServices as experimental dataset to do test. It is shown in table9.

Table 9. Contents of Mobile Service Data Set

Data set classification	concrete content
Mobile commerce user data set	500, including user identity, age, gender, occupation, consumption level
Mobile network service data set	100, service attributes include service identification, service price, service type, etc..
Scene data set	5, that is, time, location, use of equipment, activities, the surrounding personnel
Mobile commerce user history behavior context data set	163402 user historical records
Mobile commerce user behavior data set	User, the service behavior matrix, which is composed of the user's behavior variables and the behavior variables that are not used by the users.

4.3. Measuring Criteria

4.3.1. Measuring Criteria for Parallel Processing. One computing advantage of Hadoop cluster in cloud environment is parallel treatment. In the paper we use the speed-up ratio of algorithm implementation and relevant speed-up ratio to measure the performance improvement in cloud environment. Speed-up ratio is an indicator for comparing the required execution time of the algorithm for single running calculation and parallel calculation. The computing way is the ratio between single running time and parallel running time.

In the meantime, in order to find out the impact of the different number of nodes in Hadoop cluster on Hadoop cluster performance, the concept of relative speed-up ratio was raised in academic circles.

$$S = \frac{T(1)}{T(N)} \qquad (5)$$

Where, the time of stand-alone processing is $T(1)$, the time of the parallel processing of the machine is $T(N)$. The ratio of the two is the acceleration ratio.

In order to further test the effect of different number of nodes on the performance of Hadoop cluster, the academic circles put forward the concept of relative speedup, its formula is:

$$S_{\text{Relative}} = \frac{T(\sin gle\_database)}{T(many\_database)} \qquad (6)$$

4.3.2. Measuring Criteria for Recommendation Result. To evaluate the performance of recommendation system, these indicators are adopted such as: recommendation precision, recommendation result coverage, recommendation validity. With regards to which measuring criteria being chosen, the application type of collaborative filtering method plays a decisive role. Here considering both the proposed algorithm and actual situation, we take mean absolute error (MAE) and $P(u)@N$ [21] as evaluation standard.

As the commonest recommendation quality evaluation method, mean absolute error MAE determines the accuracy of recommendation result by calculating the error between predicted rating and actual rating. It can intuitively represent the goodness or badness of recommendation quality. When MAE is smaller, it means the error between predicted and actual rating is smaller and thus the recommendation quality is higher. Its formula is:

$$MAE = \frac{\sum_{i-1}^{N} |pi - qi|}{N} \qquad (7)$$

4.3.3. Scalability Analysis of Recommendation System. In order to confirm further the influence of different number of DataNode in cluster on the Hadoop cluster's performance, we use Hadoop cluster to process the same batch of testing data (Moviepilot or MobileServices dataset). We change the quantity of DataNode and record operational time of each testing dataset; then with those data, we analyze how the changing number of DataNode affects the performance of Hadoop cluster. On that foundation, we can get the performance advancement of the collaborative filtering recommendation algorithm in MapReduce framework for Hadoop cluster. The test is carried out as below: initiate Hadoop cluster; record running time of collaborative filtering recommendation algorithm after MapReduce implementation every time when one DataNode is added or removed; next calculate according to relevant speed-up ratio formula in (6) to get the speed-up ratio curve of Moviepilot and MobileServices dataset respectively on different nodes in Hadoop cluster. It is shown in Figure1.

In Figure 1, two uprising curves show the speed-up ratio of Moviepilot and MobileServices dataset on different nodes in Hadoop cluster. As seen obviously, with increasing number of DataNode in Hadoop cluster, the running speed of the entire algorithm grows nearly in a linear manner.

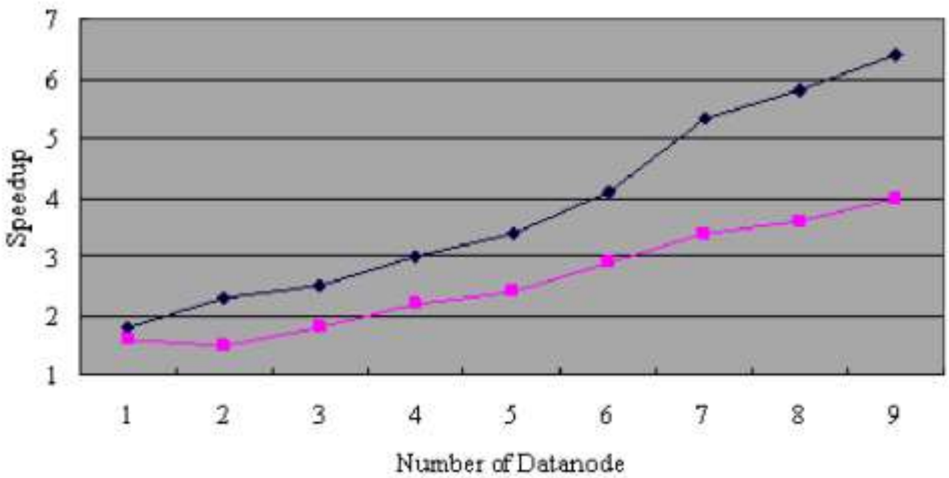

Figure 1. Performance of the Speedup of Hadoop Nodes

4.3.4 Analysis of Recommendation Result Accuracy. Mean absolute error (MAE) and $P(u)@N$ are used to evaluate the accuracy of recommendation result.

(1) Comparative testing on MobileServices dataset

When doing the comparative testing on MobileServices dataset, we use traditional situational context pre-filtering method and the improved algorithm proposed here to achieve the goal; tested dataset is user-service in MobileServices;

During testing, the period of time when target user spends is employed to pre-filter the traditional situational context pre-filtering method. The improved algorithm in the paper utilizes the collection of time which is similar to current time frame of the target user; the experimental thinking is described specifically as follows: firstly assign different values to the nearest neighbors; then observe the influence of changing number of the closest neighbors on MAE value of two recommendation approaches in any cases.

During testing, the experiment takes the way of increasing gradually the number of nearest neighbors; to be specific, assign value with 5 as a unit; choose ten testing intervals, which are 5、 10、 15、20、25、30、5、40、45 and 50, then input the traditional situational context pre-filtering recommendation method and the proposed algorithm. Finally outline different resulted MAE values and compare them. It is shown in Figure2.

(2) Comparative testing on open dataset Moviepilot

Before the comparison, firstly it classifies time stamp of training set and testing set by month; since user has only one scoring about each movie, so we replace user situation similarity with user based contextual similarity to do pre-filtering; then from training set, choose all rating records of all months which are alike the month of testing dataset as the one after pre-filtering as to do preference prediction. It is shown in Figure3-Figure4.

In the case of different number of nearest neighbors, compared with traditional situational context recommendation algorithm, the collaborative filtering recommendation method proposed here based on situation has higher precision $P(u)@N$ and lower MAE value. In short, it achieves higher recommendation quality and recommendation precision.

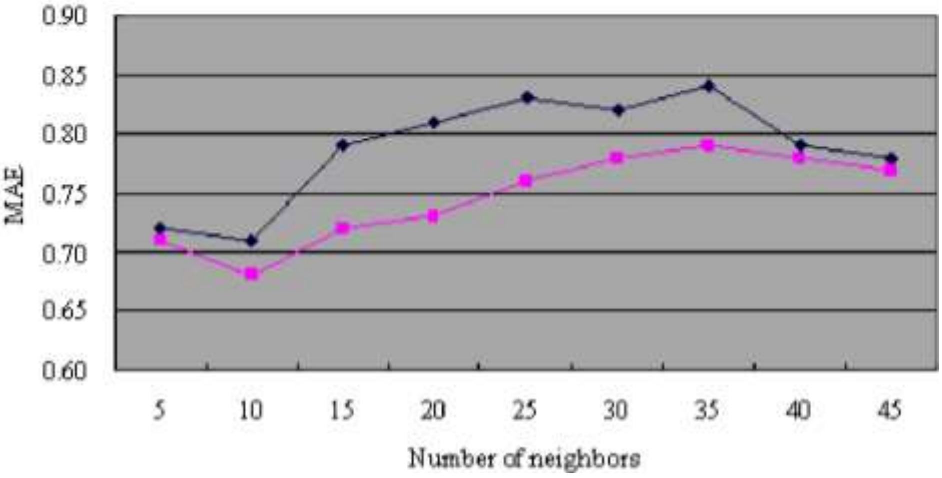

Figure 2. Comparison of MAE between Traditional Context Filtering and Proposed Algorithm

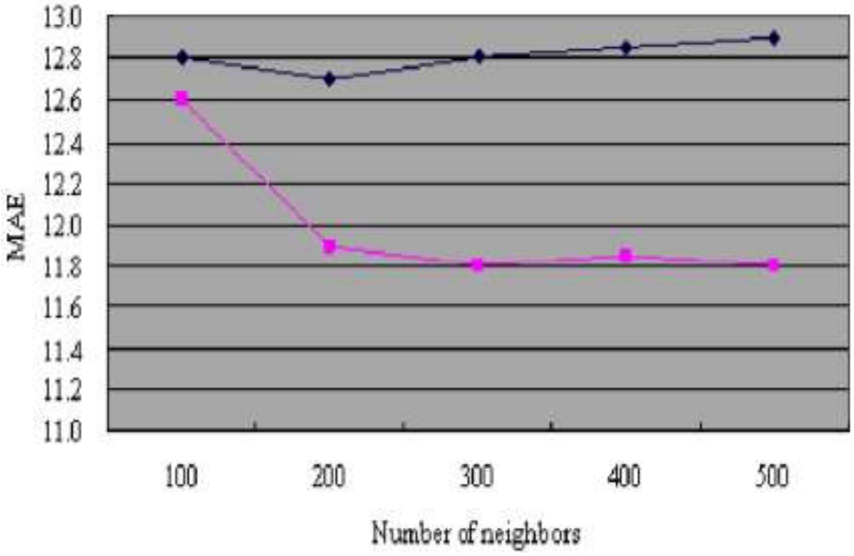

Figure 3. Comparison of MAE between Traditional CF and Proposed Algorithm

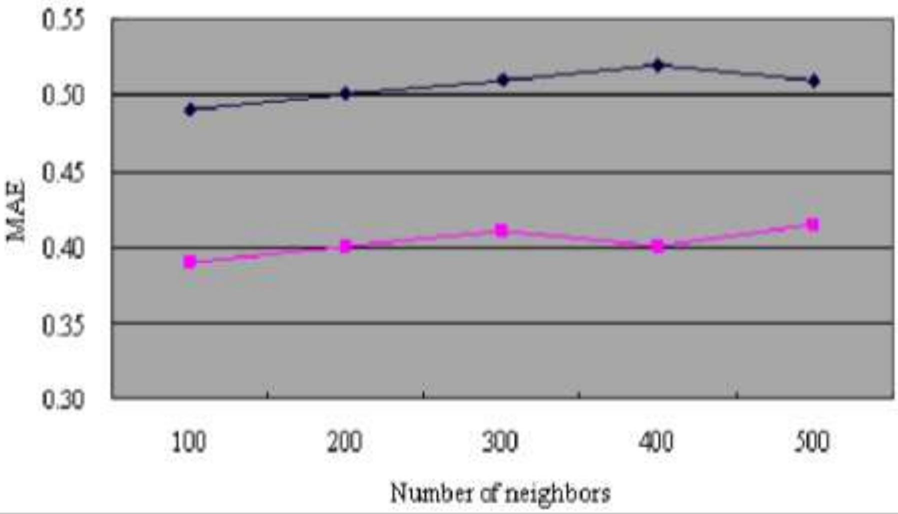

Figure 4. Comparison of $P(u)@N$ between Traditional CF and Proposed Algorithm

5 Conclusion

According to the recommendation of mobile commerce, collaborative filtering recommendation model is proposed based on the interest of mobile commerce users in collaborative filtering recommendation process.

Firstly, the model calculates the similarity of the scene based on the mobile commerce users, then, based on the similarity of the current situation of the construction of the target user set, to establish a scoring matrix based on the score of the item. Finally, a collaborative filtering recommendation method based on MapReduce is proposed. Experimental results on simulated data sets and public data sets show that compared with the traditional algorithms, in this paper, the algorithm achieves higher $P(u)@N$ accuracy and lower MAE error, so it can be used to forecast the situation of mobile commerce users in cloud environment.

Acknowledgement

This work was supported by Jilin Province Industrial technology research and development project (No. 2016C079)

References

[1] Z.-M. Chen and Z.-Q. Li, "Collaborative filtering recommendation algorithm based on user characteristics and project attributes", computer application, vol. 7, **(2011)**, pp. 1748-1750, 1755.

[2] M.-C. Jing, Zhangqin and D.-G. Tang, "Collaborative filtering system based on seed set score new user cold start recommendation", Library and information work, vol. 5, **(2013)**, pp. 124-128.

[3] X.-Y. Deng, C. Jin and Q.-P. Han, "Higuchi mouth good. Based on context clustering and user rating of collaborative filtering recommendation model", System engineering theory and practice, vol. 11, **(2013)**, pp. 2945-2953.

[4] F.-Z. Zhang, S. Liu, Z.-H. Li and J.-H. Sun, "A collaborative filtering recommendation algorithm based on the fusion of user reviews and environmental information", mini computer system, vol. 2, **(2014)**, pp. 228-232.

[5] H.-G. Rong, S. X. Huo, C.-H. Hu and J.-X. Mo, "User similarity based collaborative filtering recommendation algorithm", Journal on communications, vol. 2, **(2014)**, pp. 16-24.

[6] Z. Xia and A. B. Song, "Oriental, Luo Junzhou. Cloud computing environment based on collaborative filtering in personalization recommendation mechanism", Journal of computer research and development, vol. 10, **(2014)**, pp. 2255-2269.

[7] P. Shi, "Collaborative filtering recommendation algorithm based on user interest and project characteristics", Central South University, **(2012)**.

[8] X.-W. Zhao, "Collaborative filtering recommendation technology based on the change of user interest", Chongqing University, **(2012)**.

[9] P.-H. Zhang, "Research and implementation of personalized news recommendation system based on mobile Internet", Jilin University, **(2012)**.

[10] Q.-X. Shi, "Collaborative filtering and personalized recommendation algorithm based on user's interest in collaborative filtering", Tianjin Normal University, **(2013)**.

[11] H. Yang, "Collaborative filtering recommendation algorithm", Zhejiang Sci-Tech University project categories based on the preference, **(2014)**.

[12] M.-Q. Song, "Collaborative filtering and personalized recommendation based on the fusion of user attributes and interest contrast", Huazhong Normal University, **(2014)**.

[13] H. Hoon, "Collaborative filtering recommendation method", Beijing University of Posts and Telecommunications fusion mobile user socialization, **(2014)**.

[14] L. Miao, "Study on the personalized recommendation method of commodity based on situation", Dalian University of Technology, **(2015)**.

[15] X.-S. Zhang, "Collaborative filtering recommendation algorithm based on data sparse University of Science &amp", Technology China, **(2011)**.
[16] S.-D. Liu, "Research on recommendation technology of mobile social network based on location", Beijing University of Posts and Telecommunications, **(2015)**.
[17] H. Cai, Y.-B. Jia and C.-W. Huang, "Research on collaborative filtering recommendation method based on user trust model", Computer engineering and application, vol. 35, **(2010)**, pp. 148-151.
[18] S. Ting, "Research and implementation of micro-blog recommendation system based on cloud computing", Shandong Normal University, **(2015)**.
[19] L. Gu, P. Yang and J.-Z. Luo, "A broadcast storage under the network environment the UCL collaborative filtering recommendation method", Journal of computer research and development, vol. 2, **(2015)**, pp. 475-486.
[20] L.-C. Wang, "Research on some key technologies of context aware recommendation system", Beijing: Doctoral Dissertation of Beijing University of Posts and Telecommunications, **(2012)**, pp. 122-125.
[21] L.-C. Wang, X.-W. Meng and Y.-J. Zhang , "New approaches to mood-based hybrid collaborative filtering", Proceedings of the Workshop on Context-Aware Movie Recommendation at the 4th ACM Conference on Recommender System, **(2010)**; Barcelona, Spain.

Author

He Kun. He received his Ph.d from Jilin University of technology. He is a lecturer in College of computer science and technology, Changchun Normal University. His research interests include Computer graphics image processing.

