CONSTRUCTION OF ENTERPRISE ECONOMIC DECISION RECOMMENDATION SYSTEM BASED ON COMBINED ASSOCIATION ANALYSIS MODEL

S.J. ZHAO†

† *Institute of Finance and Economics in Shanghai University of Finance and Economics*.
merlinabFc@yahoo.com

***Abstract*— With the development of the information age and the network economy, enterprises need more and more decisions, and the difficulty and complexity of decision-making are constantly improving. The traditional centralized decision-making is no longer in line with the requirements of current social and economic development. Therefore, based on the in-depth study of conventional algorithms, this paper constructs a portfolio association analysis model for enterprise economic decision-making recommendation system, and uses the off-line test method to test the construction model. The test results show that the accuracy and recall rate calculated by the combination algorithm are higher than the common collaborative filtering algorithm. The enterprise economic decision recommendation system realizes the integration of economic mathematical model, operational research method and decision-making means, greatly improving the enterprise. The accuracy of scientifically feasible decision-making has improved the efficiency of business operations to a certain extent.**

***Keywords*— Enterprise economic decision-making; Recommendation system; Combinational association analysis model.**

I. INTRODUCTION

Through the analysis of historical modelling of the users, the system actively recommends to users to meet their interests and needs of information to help users make choices. Collaborative filtering is regarded as the most prevailing techniques for recommendation system. Slope one is a family of algorithms used for collaborative filtering. It is the simplest form of non-trivial item-based collaborative filtering based on ratings. It means that they cannot catch users different attitudes at different time (Wu, 2015). However, existing recommendation systems have common problems such as cold start, data sparsity and accuracy, which will greatly reduce the quality of the recommended system. When GeForce256 was issued in 1999, NVIDIA first proposes the concept of GPU (Graphics Processor Unit), and then a large number of complex application requirements make the whole industry flourishing.

This paper first presents the development of GPU, and then briefly introduces and compares two popular development platform of GPU: CUDA (Compute Unified Device Architecture) C and OpenCL (Open Computing Language), and finally summarizes the successful applications of GPU computing in astronomy. And in the conclusions, it also presents that the problems about branch prediction capability, the larger cache and shared memory and the thread particle size may be solved in the further research about GPU computing (Wu *et al*., 2018). The proposed method is verified by the implementation of both base and integration classifiers. Experimental results indicate that our proposed method has higher classification accuracy and faster grouping ability compared with conventional classifiers (Ke *et al*., 2018).

II. STATE OF THE ART

A prior for natural-sized LECs reduces the possibility of overfitting, and leads to a consistent accounting of different sources of uncertainty (Wu *et al.*, 2018b). A set of diagnostic tools is developed that analyzes the fit and ensures that the priors do not bias the EFT parameter estimation (Wu *et al.*, 2018c). The procedures are illustrated using representative model problems, including the extraction of LECs for the nucleon-mass expansion in SU(2) chiral perturbation theory from synthetic lattice data (Qin *et al.,* 2018). The recommendation system has been developed over 20 years. The earliest recommended system prototype was created in 1992. Goldberg et al. created a Tapestry system that could search and filter articles based on the interests and hobbies of users, and it proposed the concept of “collaborative filtering” for the first time.

The United States and European countries have already started multimedia teaching, and now have combined digital virtual teaching, live classes and lectures are also started abroad, and now the prevalence is very high, has covered the entire education system. The research results have made the recommendation system and collaborative filtering technology receive extensive attention in the subsequent research on e-commerce applications (Akhmet *et al.*, 2018). The research on personalized recommendation of products has achieved excellent research results and commercial prospects. Almost all types of e-commerce sites use a variety of forms of recommendation systems.

Foreign countries, such as Amazon, eBay, domestic Taobao, Dangdang, etc., collect and mine the historical records left by users on the site. They analyze the relevance of user interests and user buying habits and actively recommend products to users and help users get

personalized product information. This personalized recommendation is to simulate the sales staff to help users purchase the desired products.

This service not only can effectively improve the cross-selling ability of the website, but also can improve and maintain the satisfaction and trust of the customers on the website through high-quality recommendation services (Tucker *et al.*, 2014).

III. METHODOLOGY

A. Introduction to Common Algorithms

Degrees and loyalty can establish a solid bond between users and e-commerce websites. According to a paper by Amazon's former scientist Greg Linden, we know that Amazon's nearly 35% of sales come from recommended algorithms (Slawiñski *et al.*, 2006). The huge commercial value has prompted the recommendation system to become a research hotspot in academic and applied fields (Wu *et al*., 2018a). Introduction to Slope one: Slope one is an Item-Based recommended algorithm proposed by Professor Daniel Lemire in 2005. It is a form-based collaborative that is simple and unique. Because the algorithm form is very simple, it can be implemented very efficiently, and the most surprising thing is that its accuracy does not decrease compared to other complicated and computationally expensive algorithms. The general model of the Slope one algorithm can be expressed in the following manner. For a given training set θ, define the user u to score uj and ui for any item of item j and i. The following method is applied to describe the average deviation of item j relative to item i:

$$dev_{j,i} = \sum_{u \in S_{j,i}(\theta)} \frac{u_j - u_i}{card(S_{j,i}(\theta))} \tag{1}$$

Therefore, we can use the following methods to make predictions:

$$P(u)_j = \frac{1}{card(R_j)} \sum_{i \in R_j} (dev_{j,i}) + u_i \tag{2}$$

However, the above algorithm has a problem that the number of evaluations is not considered in the algorithm. For example, user U evaluates item 1 and item 2 to predict the evaluation of user item 3. If there are 100 users who have scored item 1 and item 3, 1000 users have also scored on item 2 and item 3. Obviously, the weights of two ratings are different. User U's rating of item 2 can better predict its rating on item 3. So, the weighted Slope one algorithm came into being:

$$Pw(u)_j = \frac{\sum_{i \in S(u)-\{j\}} (dev_{j,i} + u_i) c_{j,i}}{\sum_{i \in S(u)-\{j\}} C_{j,i}} \tag{3}$$

Among them, $C_{j,i} = card(S_{j,i}(\theta))$.

The general idea of the Slope one algorithm is to replace the score difference between two unknown individuals with average values. From the mathematical point of view, it is to satisfy the following formula:

$$y = x + b \tag{4}$$

Among them, x stands for the rating of item A by user u, and y refers to the rating of item B by user u. The algorithm assumes that there is a linear relationship between the scores x and y for all users on two items. A large amount of user rating data over which these two items are rated is fitted to obtain the value of parameter b. When a new user u1 has already scored A, then based on the relationship between x and y, it can be predicted that user u1 scores item B as $y_1 = x_1 + b$.

Content filtering based on KNN: The Item CF algorithm updates the Item similarity table with user behaviour at regular intervals. If there is a new Item added at this time, the Item will not exist in the Item related table, so the Item CF algorithm cannot recommend the new Item. One way to solve this problem is to update the Item similarity table frequently. However, calculating the similarity of Item based on user behaviour is very time-consuming. The main reason is that the user behaviour log is very large, and if the new item cannot be presented to the user, the user cannot be acted on it, causing these Items to fail to appear in the correlation matrix. It is necessary to use the content information of the item to improve the item cold start problem based on the Item CF algorithm. The content information of the articles is various, and different articles have different content information. This article mainly focuses on the economic decision-making of enterprises. The inherent content information of economic decision-making usually includes various resources that need to be mobilized for this decision-making, for example, personnel, advertising expenses, material quantities, and so on. With this data, we can create a data model by the following formula:

$$c_j = c / (a + b - c) \tag{5}$$

Among them, a and b are the species numbers of the two communities, and c is the species number common to both communities. For this paper, a and b indicate the number of two Item property values, and c indicates the same property value. In this way, the overall similarity score can be set by weighting.

Table 1. Similarity scores.

Attribute	Same time value	Differential value	Weight	Weighted distribution
Number of personnel	[0,1]	0	25	[0,25]
Advertising expenses	1	0	10	[0,10]
Material consumption	1	0	15	[0,15]
Others	[0,1]	0	25	[0,25]

Adding the results of the weighted similarity scores of all attributes should be distributed over [0,100]. The higher the score, the higher the similarity between the content be. Based on the calculation results of the similarity between the above contents, applying the principle of KNN can achieve related content recommendation. KNN (K-Nearest Neighbour algorithm), the K Nearest Neighbour algorithm, is a theoretically mature method.

Also, it is one of the simplest machine learning algorithms. The idea of this method is: if there is a sample in the feature space, most of the k most similar samples (nearest neighbours in the feature space) belong to a certain class, then the sample also belongs to this class. By calculating the distance or similarity between sample individuals, find the K individuals closest to each sample entity. The time complexity of the algorithm is directly related to the number of samples, and it is necessary to complete the process of pairwise comparison.

The KNN algorithm is generally used for classification algorithms. It is a supervised learning method to classify the total samples based on the training set which is given the classification rules. Here we do not use KNN to implement classification. We use KNN's most primitive algorithm idea, which is to find k content for each content that is most similar to it and recommend it to users. One advantage of using the inherent attribute of content is that the inherent attribute remains basically the same once it is created, the data output by the algorithm does not need to be recalculated once it is calculated. That is, for the content of the article, the data of the original content can be reused continuously after being initialized for the first time. It is only necessary to update the data of the newly added content. The statistical calculation of the data can use the form of incremental update, which can effectively reduce the calculation time.

B. Combinatorial Algorithm Recommendation System Model

Although there are many kinds of recommendation technologies nowadays, each recommendation technology has its own inherent advantages and disadvantages. Collaborative filtering is a widely used recommendation method, but it faces several problems. They are called cold start and data sparsity issues. Collaborative filtering requires users to jointly score at least two products.

When user evaluation data is sparse, or new users are added to the system, the quality of recommendations will be greatly reduced. The content-based recommendation technology is based on the content information of the articles and ignores the user's own preferences. In addition, based on the content recommendation, there is a great limitation on the recommendation object, and only some textual objects can be recommended. In the face of the Internet, the amount of multimedia data will greatly increase the difficulty of feature extraction.

Therefore, in order to combine the advantages of various recommendations and make up for their shortcomings, the combined recommendation technology has been adopted so far. The combined recommendation system enhances the recommendation quality by combining two or more than two recommendation technologies so as to avoid the disadvantages of each recommendation technology. Based on this, researchers have proposed the following combinations:

Table 2. Commonly used combination recommendation.

Mixed method	Description
Weighted	The recommended results generated by various recommendation technologies are selected according to the specific proportion of the system, and the results are presented to the users.
Switching	According to the actual situation of the system data change, the recommendation algorithm recommended to the user is changed. There are multiple recommendation technologies in the recommendation system, but only one recommendation technology can be used at the same time, which is decided by the specific environment.
Mixed	Give the results of different recommended methods to users at the same time
Feature Combination	Bring data features from different recommended data sources together and used it by a recommendation algorithm.
Cascade	First use one of the recommended techniques to produce a rough recommendation, and then use the second recommended technique to further produce more accurate recommendations on the previous recommendations.
Feature Augmentation	The information generated by one technology is input as the data feature of another recommendation technology, and the information generated by the former technology is used to generate recommendation to the user.
Meta-level	A model generated by a recommendation method is used as another recommended input, and the entire information generated by the former recommendation technique is used as the input of the latter recommendation technique.

The recommendation algorithm is the core of the entire recommendation system, and its merits directly determine the performance of the recommendation system to a great extent. By studying the algorithm introduced in the previous chapter, we can learn that each independent recommendation algorithm has its own defects. In the practical application recommendation system, several different recommendation algorithms are generally combined to form a combination recommendation algorithm to improve the accuracy of the recommendation.

This paper combines the improved slope one algorithm with content-based filtering to overcome the inherent flaws of the two methods and improve the recommendation quality and recommendation efficiency of the recommendation system. In Fig. 1, it is seen the structure of a common portfolio recommendation system:

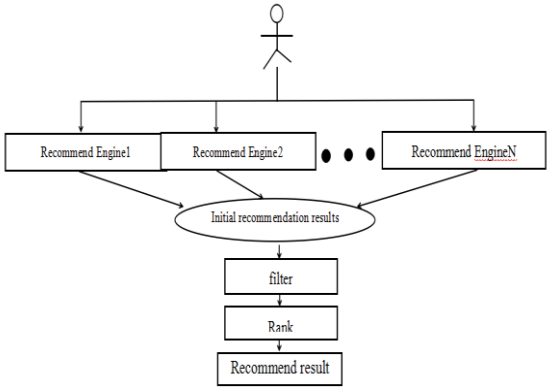

Figure 1. Architecture of recommended system.

C. The Basic Idea and Framework of Combination Recommendation Algorithm

According to the development of the current recommendation system, currently the most widely used and the most successful recommendation system is Item-based recommendation algorithm. In order to bring the advantages of this technology to full play, this paper selects the Slope one algorithm, which is simple in form, efficient in performance, and compared to other complex and costly algorithms, its accuracy is high, and the algorithm is improved. However, this algorithm still has problems such as cold start, so we consider combining a content-based filtering algorithm with it. The biggest algorithm is that it does not require the user's rating data, and its similarity calculation depends on the content of the product.

Therefore, this method can solve the cold start problem of new items, and the algorithm does not need to rely on the scoring matrix. The algorithm is also not limited by data sparsity. It can be seen that the content-based recommendation can just make up for the defects based on collaborative filtering algorithms. Item-based collaborative filtering technology does not need to consider some of the problems caused by the information of the items themselves, and it also just solves the defects of content-based recommendation.

At present, there are roughly three types of combination strategies that are suitable for this paper. One is a continuous combination of mixed recommendations. This method uses a content-based recommendation algorithm to fill up the score-predicted values missing from the user-score matrix. Then, it uses the improved Slope one algorithm to score the score matrix. Another method is to use a linear combination type of combination recommendation. This method runs two algorithms in parallel.

Then it assigns certain weights to the results generated by the two algorithms and recombines the result sets to obtain the final recommendation. The third method is the conversion type. This method is less complex than other algorithms and it is an algorithm similar to conditional checking. When the condition input by the system satisfies condition A, the result set under condition A is output. If the input of the system satisfies condition B, then the result under condition B is output for a long time.

In the context of this paper, Method 1 and Method 2 are more complex to implement. The scoring results generated by different algorithms are difficult to be unified, and there is no corresponding basis for the setting of some weight values, and it is difficult to explain the results of setting the weight values. Therefore, this paper considers the use of the third combination of strategies, which is called conversion type. For the content-based filtering algorithm, it focuses only on the user's content attributes.

The more value of the available content attributes, the higher the precision of its prediction, and it is insensitive to the data in the user-score matrix. It does not care about the user's rating and rating data. The improved Slope one algorithm, although the algorithm itself is very sparse in the user-score matrix, it can also recommend for the user, but the accuracy of the algorithm will be greatly reduced. Moreover, the Slope one algorithm cannot do anything about the cold start of new items.

To sum up, this paper considers the use of a conversion-based combination strategy. Content-based filtering focuses on the content of an item, ignoring the user's personalized needs, and fails to achieve personalized recommendation. Therefore, it cannot be used alone. The improved Slope one algorithm cannot solve data sparseness problems and item cold start problems. Therefore, the following combination strategy is used: Content filtering algorithm is used to solve the problem of data sparsity and item cold start that cannot be solved in Slope one algorithm, which is mainly based on the improved Slope one algorithm. The improved Slope One algorithm can provide personalized recommendations for users and solve the lack of content filtering based on the user's history rating. The conversion condition is the number M of user ratings. If the number of user ratings is greater than M, the system is converted to use the improved Slope one algorithm for recommendation. If the number of user ratings is less than or equal to M, the respective advantages of the two types of algorithms can be exerted to the greatest extent, and their respective disadvantages are avoided.

Using the combination recommendation technique can solve the data sparsity problem and the cold start problem mentioned above. And personalized recommendation algorithm depends on the data of user behaviours and related information of the item. How to access these data is the primary problem that the recommendation system needs to solve, and these data are often stored in the service database. Extract the relevant data files required by the recommendation system from the service database, and then pre-process the data to obtain the data required by the algorithm, then run the respective algorithms to obtain the recommended result set for each algorithm and proceed.

Then combine them to get the final result set. Based on the above considerations, this paper adopts a conversion-based combination strategy to convert the algorithm under different conditions. The specific recommended steps are as follows:

First, by pre-processing the data in the database, the data required by the KNN content filtering and the improved Slope one recommendation algorithm are respectively obtained.

Secondly, the user's identity is determined and the user's rating is used as a condition for the selection of the recommendation algorithm. For new users, recommend hot products for them. For scores less than M, run KNN-based content filtering algorithm, find nearest neighbours based on similarity, and generate Top-N list.

Thirdly, for those users whose score is greater than M, run the improved Slope one algorithm so that the similarity between articles could be calculated, and perform a score forecast on the articles based on the similarity value, and generate a Top-N list based on the score results.

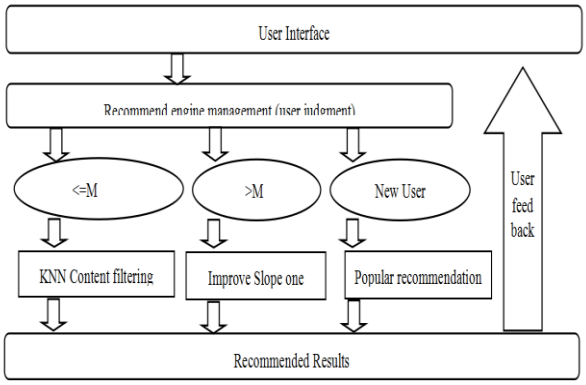

Figure 2. Model diagram of the recommended system.

IV. RESULT ANALYSIS AND DISCUSSION

How to evaluate the quality of a recommendation system is the most important issue that needs to be solved in the evaluation. A complete recommendation system typically has 3 participants: users, item providers, and websites. There are three main test methods for commonly used recommendation systems: offline experiments, user surveys, and online experiments. Offline experiments are done on the dataset, which means it does not need an actual system for it to experiment with. Therefore, one of its benefits is that the experiment does not need to be really user-involved, and it can be calculated directly and quickly so that a large number of different algorithms can be easily and quickly tested.

The user survey is performed by some real users on some tasks in the test recommendation system. When the tasks are completed, the performance of the system is understood by analysing their behaviours. Its advantage is that it can obtain the subjective feelings of the user's indicators, because the good prediction accuracy and user satisfaction are different concepts, high accuracy of prediction does not mean high user satisfaction. However, because this program requires a large number of real users to participate in, it is difficult to implement this article without considering using this program. As for the online experiment, it is also due to the fact that the existing resources and conditions are more difficult to implement, so it is not considered. Therefore, this paper will use offline testing methods to evaluate the algorithm.

The purpose of the experiment is to compare the recommendation effect of the combination recommendation algorithm and the traditional article-based collaborative filtering algorithm through experiments, and to confirm that the combination recommendation algorithm is more effective by comparing the experimental results.

The experimental method is: use the enterprise economic decision information data set to compare, use the combination recommendation algorithm and the traditional collaborative filtering algorithm to compare on different data sets, and get the experimental results.

Metrics: There are many evaluation indicators for evaluating the quality of recommendation systems, and there is no standard evaluation system. Among them, prediction accuracy is an important index to measure the ability of the recommendation algorithm to predict user behaviour. It is the most important evaluation index in off-line assessment.

Since the birth of the recommendation system, nearly 99 % of the related research papers have discussed this indicator. This is mainly because the indicator can be calculated offline, which is convenient for researchers in many academic circles to conduct algorithm research.

An offline dataset is usually needed to calculate this indicator. This dataset contains the users historical behaviour record. Then the data set is divided into training set and test set. Finally, the user behaviour model is established on the training set to predict the behaviour of the user on the test set, and the coincidence degree between the predicted behaviour and the actual behaviour on the test set is calculated as the prediction accuracy. Among them, the accuracy rate is defined as:

$$\text{Accuracy}=\frac{\sum_{\text{ueU}}|R\cap T|}{\sum_{\text{ueU}}|R|} \tag{6}$$

The recall rate is defined as:

$$\text{Recall Rate}=\frac{\sum_{\text{ueU}}|R\cap T|}{\sum_{\text{ueU}}|R|} \tag{7}$$

Under the same experimental conditions, the combined recommendation algorithm and the accuracy rate and recall rate on the two data sets. The experimental results are shown in Fig. 3:

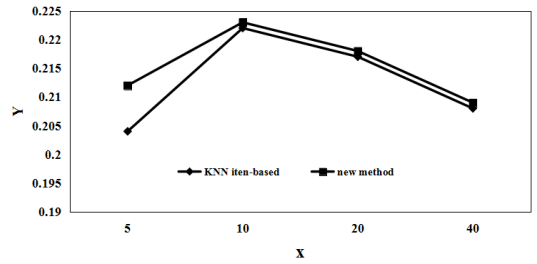

Figure 3. Data set accuracy (1).

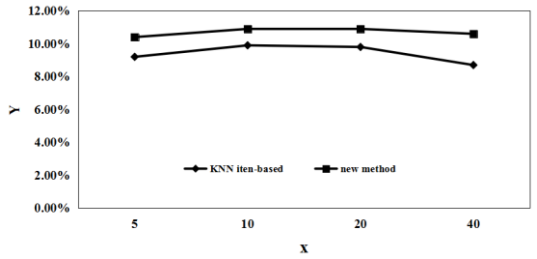

Figure 4. Data set accuracy (2).

The following uses the book rating information on an e-commerce website as a data set for experiments. Assuming that the N value is 10, the results of the obtained accuracy rate and recall rate are shown in the following table:

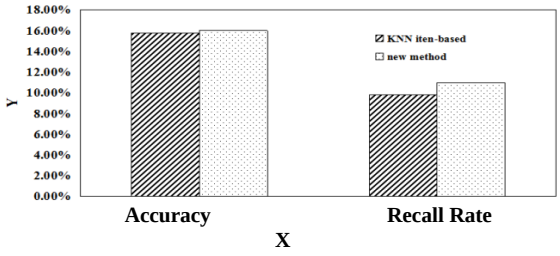

Figure 5. Comparison of accuracy rating data sets.

Through the above experimental results, we can see that the improved combined algorithm can increase the accuracy and recall to some extent. The higher the accuracy and recall, the better the performance of the algorithm be.

V. CONCLUSIONS

This paper mainly compares the performance of the new algorithm and the traditional collaborative filtering algorithm. Through simulation experiments on two different data sets, the experimental results show that the accuracy rate and recall rate calculated by using the combination algorithm are improved compared with common collaborative filtering algorithms. The reason is that when the user's rating data is very small, the new algorithm uses content-based filtering to recommend it. Although the improved Slope one algorithm can still recommend for users when the user scores less, the accuracy of the recommendation is greatly reduced. The new algorithm reduces the impact of this part of the data by combining content-based filtering, thus the performance of the algorithm has been improved to some extent.

REFERENCES

Akhmet, M. and K. Vajravelu, "Homoclinic and Heteroclinic Motions in Economic Models with Exogenous Shocks", *Applied Mathematics and Nonlinear Sciences*, **1(1)**, 1-10 (2018).

Ke, Q., S. Wu, M. Wang and Y. Zou, "Evaluation of Developer Efficiency Based on Improved DEA Model", *Wireless Personal Communications*, **102(4)**, 3843-3849 (2018).

Qin, Y., Y. Luo, Y. Zhao and J. Zhang, "Research on relationship between tourism income and economic growth based on meta-analysis," *Applied Mathematics and Nonlinear Sciences*, **3(1)**, 105-114 (2018).

Slawiñski, E., V.A. Mut and J.F. Postigo, "Stability of Systems with Time-Varying Delay". Latin American Applied Research, **36(1)**,41-48 (2006).

Tucker, J., S. Mukhopadhyay and H.M. Gonnermann, "Reconstructing mantle volatile contents through the veil of degassing[J]", *American Geophysical Union, Fall Meeting* (2014).

Wu, H., H. Zhang, L. Cui and X. Wang, "A Heuristic Model for Supporting Users' Decision-Making in Privacy Disclosure for Recommendation", *Security and Communication Networks*, **2018(4)**, 2790373 (2018b).

Wu, S., "A Traffic Motion Object Extraction Algorithm", *International Journal of Bifurcation and Chaos*, **25(14)**, 1540039 (2015).

Wu, S., M. Wang and Y. Zou, "Research on internet information mining based on agent algorithm", *Future Generation Computer Systems*, **86**, 598-602 (2018a).

Wu, S., M. Wang and Y. Zou, "Sewage information monitoring system based on wireless sensor", *Desalination Water Treatment*, **121**, 73-83 (2018c).

Received: December 15th 2017
Accepted: June 30th 2018
Recommended by Guest Editor
Juan Luis García Guirao