5-26-2012

A Product Recommendation Algorithm Based on Knapsack Optimization

Linqi Gao
College of Management, Tianjin Normal University, China, gaolqi@hotmail.com

A Product Recommendation Algorithm Based on Knapsack Optimization

Linqi Gao
College of Management, Tianjin Normal University, China
gaolqi@hotmail.com

Abstract: Personalized Recommender System becomes an important research field in Electronic Commerce, and the main goal of current recommendation models is provide Best-Service to users. But, from enterprise's viewpoint, the Max-Earning strategy is necessary to improve the benefit of enterprise. To solving this problem, knapsack model is applied to describe the commonly used Top-N recommend mechanism firstly. Then, the enterprise's earnings are described as a constraint in knapsack model, a product recommended algorithm is proposed at the basis of optimization of knapsack problem. Experimental results show the proposed algorithm has similar performance with CF model when earning requirement and amount of recommended products is lower. So, both user's value and enterprise's value are improved through the proposed algorithm.

Keywords: Recommendation System; Collaborative Filtering; Knapsack Problem; Best-Service strategy；Max-Earning strategy

1. INTRODUCTION

Up to the end of 2011, users of Internet shopping has achieved 194 millions in China [1]. So, Internet has become a more and more important market and gets more and more attention from enterprises and researchers.

With the development of Electronic Commerce, Personalized Recommendation System becomes an important research item [2, 3]. From estimating the requirement of customer, proves the suitable product and services for individual. Personalized Recommendation System aims to solving the customer's overloading of information, and improving the performance of e-Commerce system [4, 5]. At present, many e-Commerce systems have employed Personalized Recommender System in different style, such as eBay, Amazon, DangDang book store, and so on.

The recommender system not only help user to search information in Internet, but also enhance the user's satisfaction to the commercial activity. In general, the recommendation system has following effects in electronic commerce activity: Help user retrieval useful information; Promotion sales; Enhances loyalty of user

Personalized recommendation technology usually has three steps: user's requirement analysis, product search and product representation, to build an intelligent assistant for user in Internet.

Recognizing the customer's requirement is the basis of making better recommendation. More and more AI technology has been introduced to recognize customer's requirement. So, these recommendation algorithms are developed at the basis of information filtering. The unique goal is user's satisfactory.

Undoubtedly, enterprise should gain its earnings by providing the best service to users. The Best-Service strategy is the thought taken by current recommendation algorithms. These algorithms can are constructed form user's viewpoint.

But, from enterprise's viewpoint, the Max-Earning strategy is necessary to improve the benefit of enterprise. Let's consider the following case: when user's satisfactions about two products have not big differences, enterprise can actively select one product with larger earning to recommend to user. So, this recommended product will not decrease user's satisfactions and improve enterprise's earnings.

In this article, knapsack model is applied to describe the commonly used Top-N recommend mechanism firstly. Then, the enterprise’s earnings are described as a constraint in knapsack model, a product recommended algorithm is proposed at the basis of optimization of knapsack problem. Furthermore, performances of proposed algorithm are investigated under different earning constraints. Experimental results show the Best-Service strategy and the Max-Earning strategy are not Completely Opposed. In the condition with suitable earning and fewer amounts of recommended products, the proposed algorithm can meet enterprise’s benefit requirement and user's service requirement. So, both user’s value and enterprise’s value are improved through the proposed algorithm.

2. RELATED WORKS

In general, recommender system is constructed on rating data matrix. At the basis of analyzing products rated by target user, forecasting the value of unrated items, then items with higher forecasting value are selected and recommended to user [6]. The rating data can be collected by directed or concealed way, such as analyzing user’s click-stream and behavior [7].

The rating data of user forms one $m\times n$ matrix, R (m, n). In the matrix, row m represents user m, column n represents item n. The element $R_{i,j}$ in ith row and jth column represents the rating data of user i about item j, like Table 1 shows.

Table 1. User Rating Data Matrix

	Item *1*	Item *2*	……	Item *n*
user *1*	$R_{1,1}$	$R_{1,2}$		$R_{1,n}$
user *2*	$R_{2,1}$	$R_{2,2}$		$R_{2,n}$
……				
user *m*	$R_{m,1}$	$R_{m,2}$		$R_{m,n}$

To gain exact and real-time recommendation, many recommending methods have been constructed based an different theory [8,9,10], such as Bayesian Network, Clustering method, Knowledge-based System, Data Mining, Genetic Algorithm, Artificial Neural Network, Bipartite Graph, Random map theory, etc.

According to the basic thought, recommender system can be classified into two kinds: Content-based System, Collaborative System[10].

Content-based system generates recommendations through analyzing the features and similarities among products, it mainly applied the column data of Table 1. But, there are some limitations in generalization and recommending to new user.

Collaborative system generates recommendations at the basis of similarity among users. Collaborative Filtering (CF) algorithm is a successful method, widely applied in many e-Commerce systems, such as recommending movies or news for user. CF algorithm evaluates current user’s near neighbors according to the product rating data of user. Through neighbors’ rating data, the current user’s evaluation for a new product can be forecasted, then, the recommendation for current user can be gained. This kind of recommender also has defaults, such as data sparsity, new user and new product problem.

The hybrid system attempts to integrate these two kinds of methods to improve recommending capability [8, 9]. So, customer’s requirement is the basis of making better recommendation. Different AI technologies were applied to improve accuracy of forecasted ratings of target user. This is the main goal of most recommendation model. Then, according to forecasted ratings, N products are selected with higher ratings and recommended to user. This is the commonly used Top-N recommendation mechanism.

3. RECOMMENDATION ALGORITHM BASED ON KNAPSACK MODEL

The recommendation model constructed from user's viewpoint can be described as a single-goal optimizing problem. The set of recommended products for user u is selected from n candidate products. Due to the product with lower rating is not selected, the recommended set certainly has the Max-sum of rating. So, the process of recommendation is an optimization process of the sum of user's rating.

3.1 Knapsack Model

In essence, recommendation model is a knapsack model. With CF model as an example, the set of recommended products is selected from candidate products. The selecting standard is to meet user's rating. In general, a threshold is used, when a product with rating is greater than threshold, the product is recommended to user. In Top-N mechanism, the recommended set has the max-sum of rating.

The typical knapsack problem can be described as following [11]: given a set of items and a knapsack, if an item is put in the knapsack, a benefit w_j can be gained. Each item will need a capacity d_j, the knapsack has max-capacity N. The optimization task is selecting a set of items with max-sum of benefit.

$$Max \quad g=\sum w_j c_j \tag{1}$$

$$s.t. \quad \sum d_j c_j \leqq N \tag{2}$$

c_j=1 或 0，indicating the item selected or not

If setting amount of recommended products as knapsack capacity constraints, user's rating as benefit, and setting the value of d_j as 1 (This is a reasonable setting because the capacity of product can be ignored in recommendation system). So, the CF model can be described as a knapsack problem.

In this research, the performance of proposed algorithm will be inspected under different value of N. user's rating w_j can obtained through analyzing average rating of neighbor's rating. It means rating forecasting method in traditional CF model is applied in this research.

The model described by equation (1) and (2) expresses the optimization model of user's rating according to the Best-Service strategy, noted as KP1. It is obvious that the proposed algorithm has same recommended result with traditional CF model when the forecasted method of user's rating is same. So, the proposed algorithm can describe all CF model applied different AI technologies.

3.2 Requirement of enterprise's earning

The requirement of enterprise's earning can be transformed as a constraint, as equation (3). Thus, the earning requirement can be involved in optimized process. Setting the earning of product c_j as e_j, the requirement of enterprise's earning as e.

$$\text{s.t.} \quad \sum e_j c_j \geqq e \tag{3}$$

So, the model described by equation (1), (2) and (3) expresses the optimized algorithm with constraint of enterprise's earning, noted as KP2.

The difference between KP1 and Kp2 is: KP1 only concern user's satisfaction; KP2 satisfies enterprise's requirement with putting user's satisfaction as the main goal. So, KP2 can meet enterprise's benefit requirement and user's service requirement. It is an algorithm according Best-Service strategy and Max-Earning strategy.

In product database, it is easy to get product's price p_j and selling profit b_j. Then, the profit ratio e_j can be computed. In this research, profit ratio is applied as the expressions of benefit. Furthermore, the average benefit e can be gained. In experiments, the performance of proposed algorithm will be inspected respectively under different earning requirements, such as 0.8e, 0.9e, e, 1.1e, 1.2e

Knapsack problem is a typical 0-1 programming problem, there are many algorithms to solve it [12]. In experimental analysis, applying dynamic programming method provide by Matlab and user's rating forecasting method in typical CF model, the performances of proposed algorithm are discussed with different amount N of recommended products and earning requirement e. The benchmark is typical CF model with distance similarity analysis.

4. EXPERIMENTS AND ANALYSIS

The goal of experiments is comparing performances between proposed model and traditional CF model. Database View-all provides some real commercial trading data. From this dataset, Laptop trading dataset is selected to design the experiment.

Mean Absolute Error (MAE) is applied to evaluate recommendation quality. Let forecasted rating set is $\{p_1, p_2, \ldots, p_n\}$, the real rating set is $\{q_1, q_2, \ldots, q_n\}$, then

$$MAE = \frac{\sum_{i=1}^{n} |p_i - q_i|}{n} \tag{4}$$

According to the recommended result of KP1 is same with CF model, the experiments are designed to analyze performance of KP2 with different conditions. Such as:

Recommender1: KP2 with 0.8e

Recommender2: KP2 with 0.9e

Recommender3: KP2 with e

Recommender4: KP2 with 1.1e

Recommender5: KP2 with 1.2e

Experimental results show in figure 1:

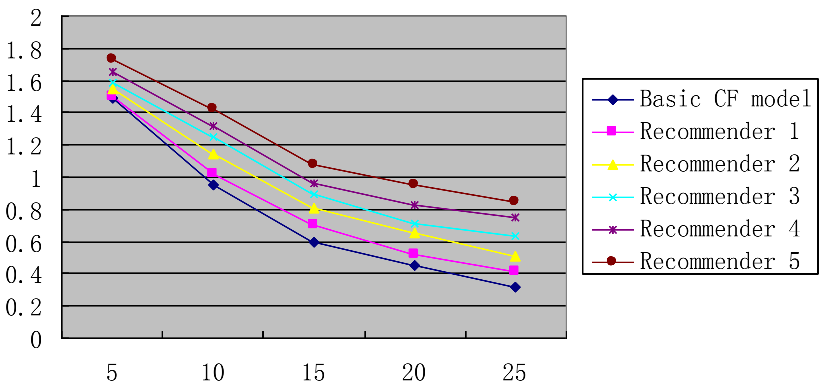

Fig. 1 Comparison among models

Through analyzing experimental results, the following conclusion can be gained:

(1) When earning requirement is lower, its MAE is lower too, closed to performance of CF model. Along with earning requirement becoming bigger, MAE become bigger too. This experimental result indicates: when earning constraint is lower, the candidate products will be more. So, recommender can select more products which satisfy user's requirement. Then, the MAE is decreased.

(2) When earning requirement is certain, the MAE will become higher according to N become larger. Its trend is similar with CF model and MAE is bigger than CF model. The reason is the proposed algorithm applies the rating forecasting method of CF model, and the additional earning requirement decreases the number of candidate products.

(3) When earning requirement and amount of recommended products is lower, performance of proposed algorithm is almost unanimously with CF model. This experimental result shows the proposed algorithm can't

decrease performance when additional earning requirement is added. It means Best-Service strategy and Max-Earning strategy can be satisfied simultaneously in this situation.

5. CONCLUSIONS

The recommendation model constructed from user's viewpoint can be described as a single-goal optimizing problem at the basis of Knapsack model. And, the process of recommendation is an optimization process of the sum of user's rating. In this optimization model, requirement of enterprise's earning can be transformed as a constraint adding into Knapsack model. So, the proposed algorithm can satisfy enterprise's requirement while putting user's satisfaction as the main goal. Experimental results show the proposed algorithm has similar performance with CF model when earning requirement and amount of recommended products is lower. That is, Best-Service strategy and Max-Earning strategy are not conflict in some situations. So, the proposed algorithm can increase enterprise's benefit and meet user's satisfactory.

In future, others dataset will be used to examine the proposed algorithm. And setting enterprise's requirement as another optimizing goal, multi-objectives optimized algorithm will be researched to simultaneously user's value and enterprise's value.

ACKNOWLEDGEMENT

This research is partly supported by National Science Fund of China under Grant No. 70402009 and Application Basis and Frontier Technology Research Fund of Tianjin, China, under Grant No.10JCYBJC26600.

REFERENCES

[1] CNNIC. 29th China Internet development statistics report. In http://www.cnnic.net.cn/dtygg/dtgg/201201/t20120116_23667.htm, 2012. l

[2] B. P. S. Murthi, Sumit Sarkar. The Role of Management Sciences in Research on Personalization. Management Science, vol. 49, no. 10, pp.1344-1362, 2003.

[3] Linqi Gao, Congdong Li. Collaborative Filtering Recommendation Model Through Active Bayesian Classifier. In Proceedings of the 2006 IEEE International Conference on Information acquisition, pp.572-577, 2006.

[4] Ralph Bergmann, Padraig Cunningham. Acquiring Customer's Requirement in Electronic Commerce. Artificial Intelligence Review, vol. 18, no. 2, pp.163-169, 2002.

[5] Judy C. R. Tseng, Gwo-Jen Hwang. Development of an Automatic Customer Service System on the Internet. Electronic Commerce Research and Application, vol. 6, no.1, pp.19-28, 2007.

[6] Gediminas Adomavicius and YoungOk Kwon. New Recommendation Techniques for Multicriteria Rating Systems. IEEE Intelligent System, May/June: 48-55 2007.

[7] Linqi Gao. Production Recommender Method Based on Customer's Behavior. Computer Engineering and Application, 2005 (3): 188-190.

[8] Manuel Barrena, Elena Jurado, Pablo Marquez-Neila, Carlos Pachon. A Flexible Framework to ease Nearest Neighbor Search in Multidimensional Data Spaces. Data & Knowledge Engineering, 69: 116-136, 2010.

[9] Amir Albadvi, Mohammad Shahbazi. A Hybrid Recommendation Technique Based on Product Category Attributes. Expert Systems with Application, 36: 11480-11488, 2009.

[10] Gediminas Adomavicius, Alexander Tuzhilin. Toward the Next Generation of Recommender Syetem. IEEE Transaction on Knowledge and Data Engineering, 17(6): 734-749, 2005.

[11] A. Jaszkiewicz. On the Computational Efficiency of Multiple Objective Metaheuristics: The Knapsack Problem Case Study. European Journal of Operational Research, 158(2): 418-433, 2004.

[12] Kostas Florios, George Mavrots, Danae Diakoulaki. Solving Multiobjective, Multiconstraint Knapsack Problems Using Mathematical Programming and Evolutionary Algorithms. European Journal of Operational Research, 203: 14-21, 2010.