2016 13th Web Information Systems and Applications Conference

A Multilayer Collaborative Filtering Recommendation Method in Electricity Market

Li Tao
State Grid Electric Power Research Institute
Nari Group Corporation
Nanjing, China
taoli@sgepri.sgcc.com.cn

Ming-hai Jiao
School of Computer Science and Engineering
Northeastern University
Shenyang, China
mjiao@mail.neu.edu.cn

Yong Dai
State Grid Electric Power Research Institute
Nari Group Corporation
Nanjing, China
daiyong1@sgepri.sgcc.com.cn

Chun-cheng Gao
State Grid Electric Power Research Institute
Nari Group Corporation
Nanjing, China
gaochuncheng@sgepri.sgcc.com.cn

***Abstract*—Transaction price accurate recommendation is a hot issue for buyer and seller on bidding information services in Chinese electricity market, a novel multilayer collaborative filtering algorithm is proposed to solve the bidding prices accurate mining problem. A three-tier relationship model of user-item-attribute is described to accommodate the real electricity transaction on bidding price service mining and recommendation. The fuzzy evaluation method is presented to improve candidate items sets of similar neighbors, integrating user fuzzy preference by attributes into membership degree evaluation. And then, the similarity function is improved to determine better proportions in items Pearson coefficients. A case study is done to give an application example for electricity market. And the experiment is also implemented to prove that the model and the algorithm are efficient and robust for application value by performance of experiment results.**

Keywords—multilayer collaborative filtering; fuzzy evaluation; information service; electricity market; recommendation algorithm

I. INTRODUCTION

Due to the abundance of choice in many online platform information services, recommender systems (RS) now play an important role [1]. It is general used as a integrated module built-in many online business application systems such as TaoBao, MeiTuan, CNKI, and so on. RS also can capture the customers preference data of web activity. And then RS targets the exact products or services as recommendation for users. Nowadays, the collaborative filtering (CF) is the main recommendation method used in existing RS [2]. It is popular for using the past activities or preferences of user ratings on items, without using user or product attribute information [3]. And the content based method is another recommendation application, using user profiles or product descriptions characteristics comparing with CF. Hybrid method combines both CF based and content based methods [4], which get the robust performance for some recommenders. Furthermore, deep learning (DL) based methods become a hot research on many applications.

The research problem of recommender systems application can categorize into following classes: news recommendation, information services recommendation, goods recommendation, and e-commerce recommendation. News recommendation methods such as context trees [5] solve redundant news titles problem published as the same context, and user personality based algorithms take dynamic feedback target for the overload information [6]. Information services use individual, collaborative, and content knowledge to analyze personalized task recommendation mechanisms [7] in the crowdsourcing information systems. Goods recommendation methods make use of associate rule [8] and preference data, but the massive new users behavior related data are difficult to detect. E-commerce recommendation methods not only use the above methods, but also adopt game theory [9], synergy theory [10, 11, 12], and cognitive psychology theory [13] according to different application areas. As to the collaborative filtering algorithms, the mining methods including K-means [14], K-neighbors [15], and Unified relevance models [16] are combined to seek accurate recommendation results. We propose a collaborative filtering methods based on items fuzzy evaluation, built on the multilayer relationship model of trade in electricity market.

Along with the advancement of electric power reform in China, the power of supply and demand side deal sizes increase recent years. The bidding transactions in market are so widely organized by electricity trading center (ETC) as to reduce social cost. How to provide valuable bidding price information for both sides of transactions is a hot mining research problem. On the other hand, the trade members are so eager to obtain overall services from online transaction information that help them to analyze bidding price and power volume of business in electricity market. So the bidding prices recommendation methods in trade platform mainly focus on

The project sponsored by the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry under grant number 49-1.

978-1-5090-5437-4/16 $31.00 © 2016 IEEE
DOI 10.1109/WISA.2016.20

collaborative filtering algorithm. Besides user based, the item based and its property based on fuzzy evaluation method are researched, built on a three tier relationship model in the paper.

The rest of this paper is organized as the follows. Section II describes the multilayer collaborative filtering methods, including three tier relationship model, items fuzzy evaluation, improved Pearson coefficient. A case study of bidding price recommendation in electricity market is presented in Section III. Then the experiment performance analyses are executed in Section IV.

II. MULILAYER COLLABORATIVE FILTERING

Traditional collaborative filtering model focus on the users and items, we extend relationship of attributes. And a fuzzy evaluation method is adopted to compute the items attribute preference values to obtain top-k candidates. Then a improved Pearson coefficient is modified by items preference weight proportion.

A. Three Tier Relationship Model

Users and items are mainly two related elements of data sets in collaborative filtering model, which is described as a binary relationship typically. Users like or dislike some items, they give a score value for items as the preference degrees. So the user based and items based collaborative filtering algorithms on the model are widely researched. But there are some less accurate items recommendation problems such as sparse users data and items data. We propose a three tier relationship model to extend attributes relationship, shown as the Fig. 1.

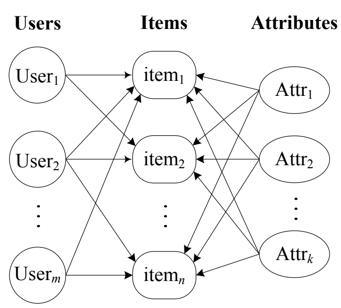

Fig. 1. Three tier relationship model.

For the recommender system, the users preferences generally rely on the items attributes. So the three tier relationship model includes ternary elements of users, items, and attributes, which relationships are described as the following: users give rates for preference items, attributes inherently affect users emotions for recommending items, and items are central bridges communicate with users and attributes.

Definition1: given the set of users $U = \{u_1, u_2, ..., u_m\}$, the set of items $I = \{i_1, i_2, ..., i_n\}$, the relationships between users and items are rating as $m \times n$ matrix.

Definition2: given the set of attributes $A = \{a_1, a_2, ..., a_k\}$, the fuzzy evaluation targets set $V = \{v_1, v_2, ..., v_p\}$, the relationships between attributes and fuzzy evaluation targets are fuzzy rating as $k \times p$ matrix.

B. Fuzzy Evaluation for Items-Attributes Perference

The preferences of items on attributes are always not clearly evaluated in real world application. And the words "good", "strong", "high" are better used to fit users emotion for items. The processes of fuzzy evaluation of items-attributes are represented as followings.

1) Elements set: The attributes set is taken as fuzzy evaluation elements set, that is $A = \{a_1, a_2, ..., a_k\}$.

2) Fuzzy evaluation targets set: The notation of targets set is marked as $V = \{v_1, v_2, ..., v_p\}$, which represents users preferences emotion such as "good", "better", "bad", and so on.

3) Single factor evaluation matrix: Each element of items has a user preference probability matrix for targets set on affected attributes, marked as preference membership degree matrix R, shown as the following:

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{k1} & r_{k2} & \cdots & r_{kp} \end{bmatrix} \quad (1)$$

The rows represent targets vectors of preferences, and columns represent attributes vectors. The element r_{ij} is a probability value given by emotion context in recommender systems.

$$\sum_{j=1}^{p} r_{ij} = 1 \quad (i=1, 2, \ldots, k) \quad (2)$$

4) Weight vector of attributes set: The attributes on items set are classic characteristics. The weight vector is a value of proportion in interval [0, 1], marked as $W = (w_1, w_2, \ldots, w_k)$. For the weight vector w_i, the following formula is met.

$$\sum_{i=1}^{k} w_i = 1 \quad (3)$$

5) Synthetic evaluation: Fuzzy operator "·" is adopted as inner multiplication about verctor W and matrix R. The evaluation formula is shown as the following:

$$B = W \cdot R \quad (4)$$

The synthetic fuzzy evaluation vector values are still between 0 and 1 due to weight interval definition.

6) Maximal membership degree principle: Each of items has p fuzzy membership values, the maximum is taken as preference evaluation value for generating top-K nearest nieghbor candidate items .

C. Improved Pearson Coefficient

The similarity is used to define the correlation degree of related items for the items based collaborative filtering.

Pearson coefficient is an excellent similarity evaluation method on items. When the sparse data of items evaluation given by users is met by above method, Pearson coefficient value is not accurate for real world yet. So we propose a improved Pearson coefficient definition, considering fuzzy membership degree proportion value of item i and item j, which includes inherent preference coefficient rating due to attributes. And improved coefficients do not appear negative value, which ranges are in [0, 1]. All are shown as follows.

$$Sim_{ij} = \frac{\min(b_i, b_j)}{\max(b_i, b_j)} \cdot \frac{\left|\sum_{i,j\in U}(r_i - \overline{r_i})(r_j - \overline{r_j})\right|}{\sqrt{\sum_{i\in U}(r_i - \overline{r_i})^2}\sqrt{\sum_{j\in U}(r_j - \overline{r_j})^2}} \tag{5}$$

The $\min(b_i, b_j)$ is the minimum of membership degree between item i and item j. The $\max(b_i, b_j)$ is the maximum of membership degree between item i and item j. The r_i and r_j are evaluation values for item i and item j given by users. And the $\bar{r}_i$ and $\bar{r}_j$ are the average scores for item i and item j given by users.

D. Multilayer Collaborative Filtering Algorithm

Multilayer collaborative filtering algorithm is still an items based collaborative filtering algorithm, but items nearest neighbors are achieved by attributes fuzzy evaluation.

1) Items nearest neighbor candidate set: The attributes set is evaluated by fuzzy membership degree method. And the set of items candidates is created by membership degree rank according to a defined threshold value θ. Then the top-K items of more than threshold values are selected as the candidate set.

2) Recommendation: Each of items in candidate set has a preference score given by users from top-K items. The items recommendation is just the nearest score prediction for related items evaluation according to weighted similarity [17] of item i and item j, shown as follows.

$$s_i = \overline{s_i} + \frac{\sum_{i,j\in I_u} Sim_{ij} \times (s_j - \overline{s_j})}{\sum_{i,j\in I_u} Sim_{ij}} \tag{6}$$

The I_u is nearest neighbors collection of items evaluated by users. The Sim_{ij} is the similarity rating of item i and item j. The $\bar{s}_i$ and $\bar{s}_j$ are the average score of item i and item j. And then the s_j is the score of item j given by users.

*3) **Algorithm**1:* Multilayer collaborative filtering (MCF) of items based algorithm is listed as follows.

Input: attributes set of items A, fuzzy evaluation matrix R for items based on attributes, target set V, weight vector W of attributes, user-item rating matrix S, the users U, the items I

Output: top-N recommendation set of items

*a) Step*1*:* Attributes set A, fuzzy evaluation targets V, and weight vector W are defined to initialize data.

*b) Step*2*:* Synthetic evaluation B for all element of items by (4) is computed to get preference membership degree.

*c) Step*3*:* A top-K nearest nieghbor candidate items is generated according to threshold value.

*d) Step*4*:* The similarity between item i and item j by (5) is computed from all candidate set according to user-item rating matrix M.

*e) Step*5*:* The prediction score for related nearest items rating by (6) is computed to get sorted top items.

*f) Step*6*:* The top N (N < K) items collection is resulted as users recommendation.

III. A Case Study for Bidding Price Recommendation

Bidding price is related to both sides of buyers and generator units in electricity market. The latter has inherent attributes which affect users preference for bidding price. So the buyers, generator units, and attributes relationship is just used for above multilayer model. We give 50 groups of units and 200 groups of electricity users data for case study to achieve the recommendation method.

A. Fuzzy Evaluation on Items

The fuzzy evaluation on items in electricity market is implemented for generator units evaluation on attributes, which are set as A = {soot, sulfur dioxide, oxynitride, sewage, capacity}. The target elements are set as V = {excellent, good, bad}.

1) Membership degree evaluation

a) Single item fuzzy evaluation matrix: The single item is about the generator unit i with attributes set A. And the evaluation matrix is composed of target set V preference probability. For example, the attributes element "soot" is rated as "excellent" for *unit*1, which probability is 0.35. The probability of "good" and "bad" is as 0.2 and 0.45 respectively. The unit of fuzzy evaluation matrix R_{unit1} is as follows.

$$R_{unit1} = \begin{bmatrix} 0.35 & 0.20 & 0.45 \\ 0.55 & 0.10 & 0.35 \\ 0.60 & 0.25 & 0.15 \\ 0.50 & 0.25 & 0.25 \\ 0.72 & 0.18 & 0.10 \end{bmatrix}$$

b) Weight vector of attributes set: For the attributes A of *unit*1, the weight vector is determined as W = (0.05, 0.2, 0.2, 0.05, 0.5).

c) Synthetic evaluation: The fuzzy evaluation of *unit*1 is computed as B = (0.6325, 0.1825, 0.185) according to (3). The maximal membership degree value is 0.6325.

2) Top-K nearest neighbor candidate items

All the 50 groups of items are evaluated to achieve the attributes targets "excellent" preference membership degree, shown in Fig 1. We give a threshold value 0.4 for remaining nearest neighbor candidate items. The 30 groups of items are selected as candidate items for collaborative filtering.

B. Items Based Collaborative Filtering

As to the sparse data of users scores for items, that is, the invalid users are filtered out. And the 126 groups of users scores are used for items based collaborative filtering, which proportion is corresponding to the candidate items proportion.

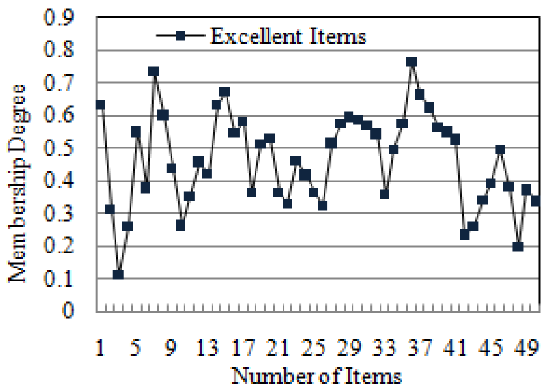

Fig. 2. Membership degree of 50 items.

All electricity users scores for item unit *i* and *j* are listed to compute generator unit similarity of Pearson coefficient. Considered the proportion between minimum and maximum of membership degree, the improved Pearson coefficient is more stable than traditional Pearson coefficient method.

Recommendation process is advanced to obtain the score prediction of items neighbors, that is, each generator unit is predictably scored by nearest neighbors of generator unit. The score results are obviously stable for top N recommendation collection of generator units. Because each generator unit give a bidding price in electricity market, the recommended item *i* and *j* are consistent with the score respectively. Sorted score is resulted in bidding price offered by generator unit in the State Grid. The recommended price of item is better accurate than traditional recommendation method, shown as Fig. 3.

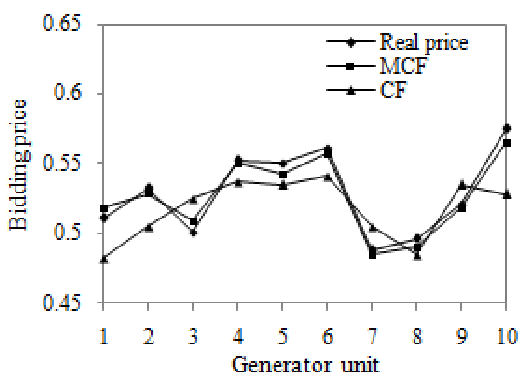

Fig. 3. Price comparison with two methods.

The traditional items based collaborative filtering (CF) algorithm is used to bidding price recommendation with a smooth curve. The proposed MCF algorithm is more accurate recommending price result than CF algorithm with price result referred to real users history bidding price. Meanwhile, the price unit is Yuan per KWh in Chinese electricity trade market.

IV. Experiments and Analyses

A. Dataset and Evaluation Metrics

The dataset of MovieLens [18] is selected in order to test algorithm performance with experiment. The 100,000 ratings and 6,100 tag applications applied to 10,000 movies by 700 users. The tags are the attributes related for movies items. The ratings are scored by users with 5 star mark of integer 1, 2, 3, 4, and 5, which integer 5 means “excellent” and integer 1 means “strong bad”. So the evaluation targets are including “excellent”, “perfect”, “good”, “bad”, “strong bad”. All the movies items are processed by fuzzy evaluation related to tags attributes. The dataset is randomly sampled by 75% as training dataset, other 25% dataset is used as testing dataset.

The prediction accuracy is the main indicator employed to evaluate collaborative filtering algorithm. And the mean average error (MAE) [19] and average accuracy rate (AAR) are two key evaluation metrics. For the recommending top N items, the mean average error is calculated the deviation between the predicted ratings p_i and the items real ratings r_i by users scores. The smaller the MAE value is, the better the recommendation accuracy performance is. The MAE is defined as follows.

$$MAE = \frac{\sum_{i=1}^{N} |p_i - r_i|}{N} \tag{6}$$

Another metric of average accuracy rate is the precision for items recommendation, which is popular in the information retrieval and data mining field. The definition is the proportion with correct recommendation numbers N_r of items and all recommendation numbers N of items, which include correct and wrong items in RS. Then the precision is defined as follows.

$$\Pr ecision = \frac{N_r}{N} \tag{7}$$

B. Experiment Results and Performance Analyses

The experiments data are computed to compare with our proposed MCF algorithm, traditional CF algorithm, and CBES algorithm in [20] according to the evaluation metrics.

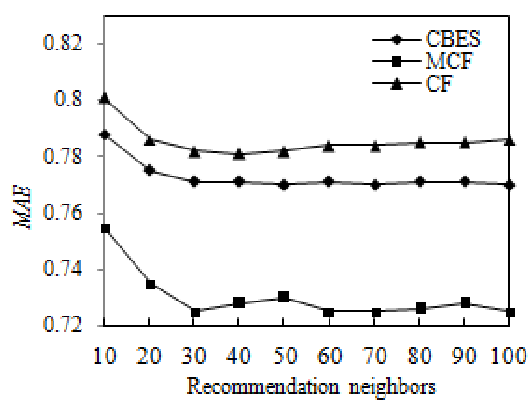

Fig. 4. Algrothms comparison on MAE.

The MCF algorithm is mainly contributed to get a top K candidate items at the fuzzy evaluation stage. Then the top N items are filtered by collaborative similarity at the second stage. So the recommendation precision is obviously improved by MCF algorithm contrast with CF algorithm and CBES algorithm. Also is the MAE, shown as Fig. 4 and Fig. 5.

As to the comparison on MAE, the horizon axis is the numbers of recommendation neighbors, which scale is increased to 100 by recommendation 10 clusters. And the threshold θ value of fuzzy evaluation is set as 0.4. The MAE values show 7.5% lower by MCF algorithm than traditional CF items based algorithm.

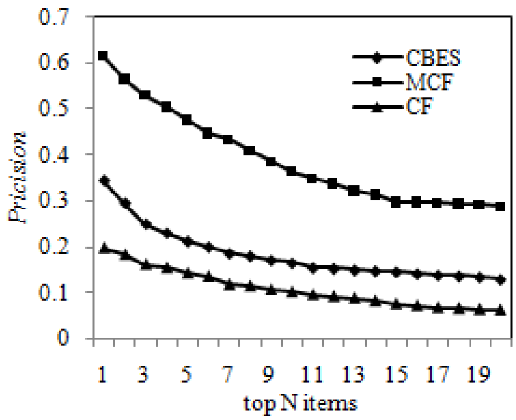

Fig. 5. Algrothms comparison on precision.

As to the comparison on precision, the horizon axis is the numbers of top N items neighbors, which scale is increased to 20 by one unit interval. All three algorithms have a monotone decreasing trend of precision curve. And the precision is stable at the 15 items neighbors. The precision values also show 7.6% higher by MFC algorithm than CF algorithm.

V. Conclusion

The multilayer collaborative filtering algorithm is proposed in this paper to solve the accurate mining problem. The fuzzy evaluation between items and attributes is presented to improve candidate items sets. And the similarity function is improved to determine better proportions in items Pearson coefficients. The experiment is implemented to give the better algorithm results and analyses, comparing MCF algorithm with other algorithms.

References

[1] W. Zhang, H. Sun, X. Liu, and X. Guo. "Temporal qos-aware web service recommendation via non-negative tensor factorization," Proceedings of the 23rd International Conference on World Wide Web, New York, pp. 585-596, April 2014.

[2] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez. "Recommender systems survey," Knowledge Based Systems, Amsterdam, vol. 46, pp. 109-132, July 2013.

[3] H. Wang, N. Wang, D. Yeung. "Collaborative deep learning for recommender systems," Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, pp. 1235-1244, August 2015.

[4] L. Hu, J. Cao, G. Xu, L. Cao, Z. Gu, and C. Zhu. "Personalized recommendation via cross-domain triadic factorization," Proceedings of the 22rd International Conference on World Wide Web, Rio De Janeiro, Brazil, pp. 595-606, May 2013.

[5] F. Garcin, C. Dimitrakakis, B. Faltings. "Personalized news recommendation with context trees," Proceedings of the 7th ACM Conference on Recommender Systems, Hong Kong, pp. 105-112, October 2013.

[6] S. Dooms. "Dynamic generation of personalized hybrid recommender systems," Proceedings of the 7th ACM Conference on Recommender Systems, Hong Kong, pp. 443-446, October 2013.

[7] D. Geiger, M. Schader. "Personalized task recommendation in crowdsourcing information systems - Current state of the art," Decision Support Systems, Amsterdam, vol. 65, pp. 3-16, September 2014.

[8] S. Long, W. Zhu. "Mining evolving association rules for e-business recommendation," Journal of Shanghai Jiaotong University (Science), Shanghai, vol. 17, pp. 161-165, February 2012.

[9] J. Zhao, T. Zheng. "Taobao C2C mode game model analysis," Journal of Management, Xinxiang, vol. 27, pp. 58-61, June 2014.

[10] C. Miao, J. Feng, L. Sun, L. Ma. "Enterprise competence systems evolution model based on synergistic theory and self-organization theory," Journal of Nanjing University of Science and Technology, Nanjing, vol. 37, pp. 192-198, February 2013.

[11] S. Cai, Q. Yuan, P. Zhou, F. Liang. "Collaborative filtering recommendation model in micro-blogging website based on information diffusion theory," Systems Engineering-Theory & Practice, Beijing, vol. 35, pp. 1267-1275, May 2015.

[12] P. Zhang, C. Li, L. Yang. "Study on operational efficiency of electricity market based on synergetic theory," East China Electric Power, Shanghai, vol. 38, pp. 1659-1662, November 2010.

[13] H. Sun, Q. Xing. "The influences of feedback on perceptual category learning and its' cognitive and neurophysiological mechanisms," Advances in Psychological Science, Beijing, vol. 22, pp. 67-74, January 2014.

[14] L. Ungar, D. Foster, E. Andre, et al. "Clustering methods for collaborative filtering," Proceedings of the AAAI Workshop on Recommendation Systems, Madison, Wisconson, pp. 112-125, July 1998.

[15] C. Li, C. Liang, L. Ma. "A collaborative filtering recommendation algorithm based on domain nearest neighbor," Journal of Computer Research and Development, Beijing, vol. 45, pp. 1532-1538, September 2008.

[16] J. Wang, A. Vries, M. Reinders. "Unified relevance models for rating prediction in collaborative filtering," ACM Transaction on Information Systems, vol. 26, pp. 1-40, March 2008.

[17] J. Bobadilla, F. Ortega, A. Hernando, J. Bernal. "A collaborative filtering approach to mitigate the new user cold start problem," Knowledge Based Systems, Amsterdam, vol. 26, pp. 225-238, February 2012.

[18] GroupLens. "Movielens," http://grouplens.org/datasets/movielens, January 2016.

[19] L. Herlocker, J. A. Konstan, L. G. Terveen, J. T. Riedl. "Evaluating collaborative filtering recommender systems," ACM Transactions on Information Systems, New York, vol. 22, pp. 5-53, January 2004.

[20] W. Kong, Q. Liu, Z. Yang, S. Han. "Collaborative filtering algorithm incorporated with cluster-based expert selection," Journal of Information & Computational Science, Hong Kong, vol. 9, pp. 3421–3429, December 2012.