Journal of Computers Vol. 30 No. 2, 2019, pp. 14-26
doi:10.3966/199115992019043002002

The Recommendation Method of Power Selling Packages Based on the Optimal Feature Subset of Power Trading Users

Zhao-Yang Qu[1,2], Rong-Qiang Feng[1,2*], Nan Qu[3], Yao-Wei Liu[4], Jia Yan[4], Chong Qu[5]

[1] School of Information Engineering, Northeast Electric Power University, Jilin 132012, China
QZYwww@neepu.edu.cn, frqkycg@163.com

[2] Jilin Engineering Technology Research Center of Intelligent Electric Power Big Data Processing, Jilin 132012, China
QZYwww@neepu.edu.cn, frqkycg@163.com

[3] Maintenaue Company of Jiangsu Power Company, NanJing 210000, China
1406989666@qq.com

[4] State Grid Jilin Electric Power Supply Company, Jilin 130000, China
1569666840@qq.com

[5] Fushun Power Supply Company of State Grid Liaoning Electric Power Supply Company, LiaoNing 113001, China
lancie816@qq.com

Received 29 July 2017; Revised 28 November 2017; Accepted 9 January 2018

Abstract. By aiming at huge user groups in the power market and the difficult selection of power selling packages in the trading process, a recommendation method of power selling packages based on the optimal feature subset of power trading users was proposed. First of all, the Optimal Feature Subset of Power Trading Users (OFSPTU) was defined, and the optimal subset discovery algorithm based on coverage rate with weighing increase was designed, the optimal user feature subset could be screened out reasonably from mass trading users. Next, the similar computing method of power selling packages based on attribute correlation was proposed, in the form of clustering and confirming package attribute weighing, the package similarity was calculated, obtaining the nearest neighbor of package projects. At last, the accurate recommendation of power selling packages based on the OFSPTU and package attribute was realized. The experiment shows that the effectiveness of OFSPTU and accuracy of the recommended algorithm.

Keywords: attribute weighing, optimal feature subset, power market, project coverage rate, recommendation of power selling packages, weighing increase

1 Introduction

No. 9 document of the power reform of China indicates that the nation pays the high attention to reform the power selling market [1]. Power selling is the core in power consumption. As a commodity, electric power has the natural user viscidity, it can derive and increase value-added services for numerous users. In addition, these services will be constantly discovered and created, the value space is great. Power selling packages refer to services under the internet thinking and the innovation of operational mode, they are also inevitable products in the competitive electric power market. According to statistics, the power

* Corresponding Author

terminal market in Germany has already provided more than 9000 power selling packages for users since 2007 [2-3]. Texas opened the prelude of power selling market reform in 1999. After 14 years of development, 1820 power selling packages with fixed rate, pricing models and strategies could be provided on the powertochoose, for the sake of being convenient for users to select suitable power suppliers and power selling schemes [4-5]. The time-of-use power price packages are provided for users in the UK after the power reform, there are hundreds of packages [6]. The complicated quantity, categories and contents of power selling packages also bring relevant troubles to users in package matching and selection.

By virtue of collaborative filtering and content-oriented recommendation, the recommendation system can effectively provide the individualized recommendation results, which have the high correlation with interest preference of target users, so as to solve a problem of difficult user selection for users. At present, lots of scholars have studied it in some aspects. In [7-8], the recommendation was finished by measuring the relationship between users or products to look for the similar neighbor set, but the problem that unpopular projects were easy to be submerged in popular projects wasn't considered. The contribution rate of unpopular projects for recommendation was neglected. In [9], the popular bias problem of contribution rate in unpopular and popular projects of the recommendation was solved, but attribute features of projects weren't considered in the recommendation project, thus accuracy of recommendation results couldn't be guaranteed. In [10-12], project attribute features were introduced in the recommendation to solve a cold start problem, it realizes the individualized recommendation for different users and diverse preferences, but influences of huge user groups on recommendation efficiency were neglected. In [13-14], user groups were analyzed. By deleting inactive users, user scale to be analyzed was reduced. Multiple user features were comprehensively used to improve the corresponding effects. In [15], a typical user group is searched for based on data categories and user ratings. In addition, the group clustering mode was used to shorten user groups. This was a common means. In [16], it showed that proper clustering could be used to avoid data sparsity and effectively improve accuracy. In [17-18], the clustering method could be used to propose the convergent-divergent nearest neighbor range, thus the corresponding algorithm could adapt to the large-scale data set. In [19], the genetic algorithm was used to optimize initial point selection as user clustering. But the above-mentioned group processing methods might have some inherent defects: clustering process made some inactive users or projects submerge in the large categories; it couldn't ensure that user interests could be sufficiently reflected, while guaranteeing precision. As a result, it is necessary to realize the new user subset screening method, thus the screened subset not only can fully embody interest preferences of all users, but also can cover and include known projects as many as possible.

To sum up, on the basis of considering contribution rate of unpopular projects, project features and user groups, the recommendation was introduced in the field of power selling packages, a recommendation method of power selling packages was proposed based on coverage rate with weighing increase. On the one hand, the concept of OFSPTU was defined, optimal subset discovery algorithm based on coverage rate with weighing increase was proposed, hoping to screen out an optimal subset when all users conduct package transaction on the basis of the higher project coverage rate and the relatively accurate scoring. Meanwhile, the problem of unpopular projects' low exposure rate in the recommendation process was solved. On the other hand, on the basis of obtaining the optimal subset, attribute weighing of power selling packages was combined to conduct the accurate recommendation of power selling packages.

2 Relevant Definitions of Power Selling Package Recommendation

2.1 Coverage Rate of Power selling Package Projects

The set of power trading users is defined as:

$$U = \{U_i\} \qquad 0 \le i < |U| . \tag{1}$$

The set of power selling package projects is defined as:

$$P = \{p_j\} \qquad 0 \le j < |P| . \tag{2}$$

The interactive information between power trading users and power selling package projects is expressed as the scoring matrix R:

$$R = \{R_{ij}\}_{|U|\times|P|} \qquad r_{ij} \ge 0 . \tag{3}$$

$r_{ij}=0$, showing that the power trading user u_i doesn't score the power selling package project p_j. In other words, the behavior of u_i doesn't cover the power selling package project p_j. For the power trading user subset U', the coverage set of the power selling package project is defined as:

$$P_{U'} = \{P_k\} \qquad \exists r_{ik} \ne 0, u_i \in U' . \tag{4}$$

The coverage rate of power trading subset U' in the power selling package project is expressed as:

$$Cov(U') = \frac{|P_{U'}|}{|P|} \times 100\% . \tag{5}$$

For the power trading user subset with the size of k, the coverage rate of the power selling package project has the monotone increasing, namely:

$$0 \le Cov(U_1) \le Cov(U_1 \cup U_2) \le 1 . \tag{6}$$

2.2 The Coverage Rate of Power Selling Package Projects with Weighing Increase

The coverage rate of power selling package projects with weighing increase after the power trading subset U_1 joins the user subset U' is defined as:

$$ICov_{U'}(U_1) = Cov(U_1 \cup U') - Cov(U') . \tag{7}$$

The coverage set of two power trading user subsets' power selling package projects can't be confirmed as the empty intersection, so $Cov(U_1)>Cov(U_2)$ may be not established as $ICov_{U'}(U_1)>ICov_{U'}(U_2)$.

Make $U_1=\{u\}$, adding U' in an power trading user u, we can obtain:

$$ICov_{U'}(u) = \frac{|P_{u\cup U'}| - |P_{U'}|}{|P|} \times 100\% . \tag{8}$$

The same power selling package projects in P_u and $P_{U'}$ will be mutually cancelled out, thus:

$$ICov_{U'}(u) = \frac{\sum_{p_j \in P} f\left(p_j \in P_u, p_j \notin P_u\right)}{|P|} \times 100\% . \tag{9}$$

Where $f(\,)$ is

$$f\left(p_j \in P_u, p_j \notin P_u\right) = \begin{cases} 1, p_j = 1 \\ 0, p_j = 0 \end{cases} . \tag{10}$$

In the formula, $p_j=1$, showing existence of p_j; $p_j=0$, showing inexistence of p_j. Power trading users are long-tail for power selling package projects, thus unpopular power selling package projects are always covered by popular projects. In order to solve such a problem, the coverage rate of power selling package projects with weighing increase is defined as:

$$ICov_{U'}(u) = \frac{\sum_{p_j \in P} \omega_j \times f\left(p_j \in P_u, p_j \notin P_u\right)}{|P|} \times 100\% . \tag{11}$$

Where, weight ω_j is

$$\omega_j = 1 - \frac{\lg\left(S_j\right)}{\lg\left(\left|U\right|\right)}. \quad \textbf{(12)}$$

Where, S_j means that a total of S_j power trading users select the power selling package project P_j, showing that the more number of power selling package projects is selected, the higher weighing will be obtained by unpopular power selling package projects, so as to improve the possibility that unpopular power selling package projects are covered by popular projects.

2.3 Definition for the OFSPTU

The OFSPTU can stand for interest preferences of all users in the power selling package transaction as many as possible, thus the subset scoring error $Err(U')$ can be expressed as the interest feature bias between power trading user subset and all power trading users, namely:

$$Err(U') = \sum_{p_j \in P} \frac{\left|avg(p_j, U') - avg(p_j, U)\right|}{avg(p_j, U)} \times 100\%. \quad \textbf{(13)}$$

Where, $avg(P_j, U')$ is the average score of the power selling package project pj in the power trading user subset U', showing that the smaller $Err(U')$ is, the greater representative power trading user subset U' will be for interest preferences of all power trading users. However, if U' is small, it shows the power trading user subset U' can't cover all power selling package projects as many as possible.

As a result, OFSPTU is one of the subsets in the power trading user set U. It should meet the following requirements: to minimize $Err(U')$ and maximize $Cov(OFSPTU)$.

In the study, the author focused on studying how to maximize $Cov(OFSPTU)$ of OFSPTU.

3 Recommendation Process Analysis of Power Selling Packages Based on OFSPTU

In the package trading process, due to representativeness of optimal feature subset and the higher project coverage rate, it has the large advantage to use optimal feature subset to realize recommendation of power selling packages. The nearest neighbor searching range of users can be reduced to optimal feature subset from all users, so as to reduce time consumption of the power selling package recommendation algorithm. Meanwhile, it won't lose accuracy degree of the recommended results. For this reason, a recommendation algorithm of power selling packages based on OFSPTU was proposed. On the basis of considering optimal feature subset, the attribute features of clustering and power selling packages should be also brought into the recommendation, so as to reduce the searching space of power selling package projects and improve accuracy of package recommendation. The specific process is shown in Fig. 1.

The recommendation algorithm will consider from two directions (power trading users and power selling package project attribute) by combining with collaborative filtering recommendation algorithm based on projects and users. The coverage rate based on weighing increase should be considered to design the optimal subset discovery algorithm from the user perspective. The optimal feature subset was screened out from power trading users. From the perspective of power selling package projects, the similarity calculation method of power selling packages based on attribute correlation was proposed. By establishing the package attribute matrix, the k-means clustering method was used to cluster power selling package projects. Moreover, by aiming at the different attribute priority when power trading users select power selling packages, the analytic hierarchy process (AHP) was used to calculate weighing of project attribute, so as to find out the nearest neighbor project set in similarity between package projects in target power selling package projects' project clustering for initial predictive scoring. At last, the obtained predictive scoring similarity in OFSPTU was used to search for the nearest neighbor in target users. According to scoring situations of the nearest neighbor, unscored power selling package projects for target users were conducted the final prediction to select N highest scoring package projects to target users.

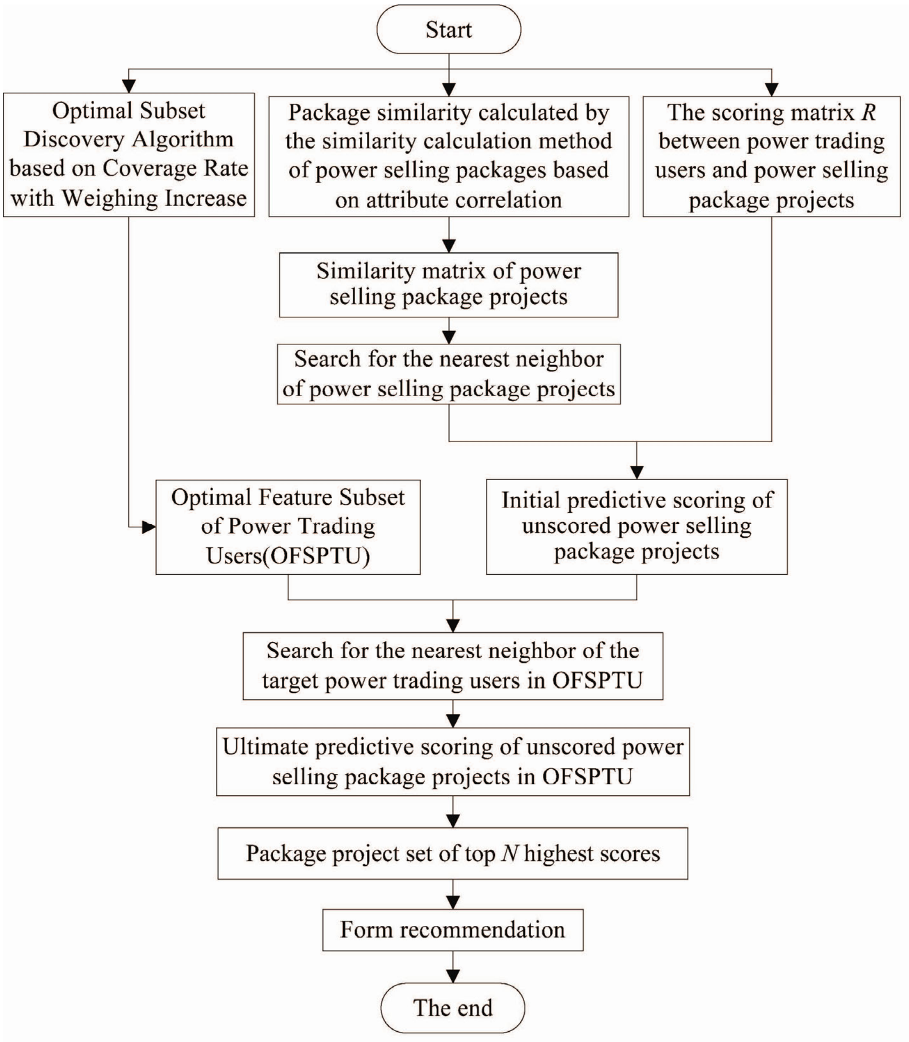

Fig. 1. Recommended flow chart of power selling package based on OFSPTU

4 Optimal Subset Discovery Algorithm Based on Coverage Rate with Weighing Increase (OSDA_CRWI)

Coverage rate of weighing increase projects can solve the defect that unpopular package projects may be submerged in popular projects. By combining with the coverage rate of weighing increase projects, the OSDA_CRWI was proposed to improve exposure rate of unpopular package projects, shorten the searching areas for power selling package recommendation and reduce calculated quantity. The specific steps are described as follows:

Input: power trading user set U, power selling package set P, scoring matrix R, and size of OFSPTU k

Output: OFSPTU

Step 1. According to the scoring matrix R between power users and power selling packages, the power trading users out of R should be removed from U. Meanwhile, power selling packages out of R should be removed from P;

Step 2. An empty power trading user subset $U'=\Phi$ should be created. Meanwhile, a candidate power trading user set $U_c=U$ should be also established.

Step 3. For any power trading user in U_c, the formula (11) should be used to calculate coverage rate of power selling package projects with weighing increase. The power trading user u with the maximal coverage rate should be added in U'. Meantime, u should be removed from U_c.

Step 4. The step 3 should be repeated until k power trading users. At last, U' should be outputted.

U' is the required OFSPTU.

5 The Similarity Calculation Method of Power Selling Packages Based on Attribute Correlation

Attribute aims at project features. For recommendation, the stability is higher and can really embody the relationship between users and projects, as well as the relationship between projects. Meanwhile, it also can solve a problem of cold start. Considering the similarity calculation methods of package attribute factors, the steps are shown as follows:

Input: the power selling package project set *P*

Output: Similarity matrix of power selling packages *W*

Step 1. According to power selling package project set *P*, the attribute matrix $A=(a_{ij})_{n\times m}$ of $n\times m$ order power selling package projects should be constructed. a_{ij} is the j^{th} attribute of the power selling package project *i*.

Step 2. The attribute value a_{ij} is expressed by the binary variable (0,1). If the power selling package project *i* includes j^{th} attribute, $a_{ij}=1$, without $a_{ij}=0$.

Step 3. Every power selling package project in the matrix *A* is considered as a vector and it is regarded as the peak in the space for package project clustering to find out the cluster of the target project.

Step 4. AHP should be used to confirm weighted value ω of every power selling package attribute in matrix *A*.

Step 5. In the target clustering, comprehensive similarity of two power selling package projects should be calculated by weighing, so as to obtain the similarity matrix *W* of power selling packages. The similarity of the package *i* and the package *j* at the k^{th} attribute can be expressed as $\omega_k/1+|a_{ik}\text{-}a_{jk}|$. The similarity between power selling package projects i and j stands for the similarity of their comprehensive attribute features. The computational formula is shown as:

$$asim(i,j)=\sum_{k=1}^{t}\frac{\omega_k}{1+\left|a_{ik}-a_{jk}\right|} \qquad 0\le\omega_k\le 1\,. \tag{14}$$

Where, ω_k is the k^{th} attribute feature weight.

6 Recommendation Algorithm Implementation of Power Selling Packages

On the basis of optimal feature subset and package attribute similarity, the recommendation method of Power Selling Packages based on OFSPTU (PSPO) was realized. The initial scoring prediction based on project attribute was calculated. The ultimate scoring prediction based on user side was calculated. Both of them complete teach other’s advantages. The specific expression is shown as follows:

Input: power trading user set *U*, similarity matrix *W* of power selling packages, package scoring set *R*, and OFSPTU

Output: Recommended power selling set *P′*

Step 1. According to the similarity matrix *W* of power selling packages, the nearest neighbor set *KNN*(*i*) of the package project *i* can be obtained.

Step 2. Based on the improvement of the collaborative filtering algorithm, the package scoring set *R* should be combined. *asim*(*i*, *j*) is considered as the score of target power selling package project for the nearest neighbor set *KNN*(*i*) of average weighing to generate the initial predictive scoring of user *u* for the target projct i. It can be expressed as:

$$P(u,i)=\overline{r_i}+\frac{\sum_{j\in KNN(i)} asim(i,j)(r_{uj}-\overline{r_j})}{\sum_{j\in KNN(i)} asim(i,j)}\,. \tag{15}$$

$\overline{r_i}$ is the average scoring of user *i* for power selling package project. r_{uj} is the scoring of user *u* for power selling package project *j*.

Step 3. The similarity of the target user *u* and every feature user *v* in OFSPTU should be calculated. The k^{th} user with the highest similarity with u should be selected from OFSPTU as the nearest neighbor set U_N. The similarity of power trading users *u* and *v* is expressed as:

$$sim(u,v)=\frac{\sum_{i\in I_{uv}}(r_{ui}-\overline{r_u})(r_{vi}-\overline{r_v})}{\sqrt{\sum_{i\in I_{uv}}(r_{ui}-\overline{r_u})^2(r_{vi}-\overline{r_v})^2}}\qquad v\in OFSPTU\ . \tag{16}$$

Where, I_{uv} stands for the project set scored by users u and v. $\overline{r_u}$ is the average scoring of the user u's evaluated power selling package projects.

Step 4. For the unscored power selling package project p of the target user u, the ultimate predictive score of u is obtained by U_N's weighing average for p. The predictive scoring formula of user u for the project is shown as follow:

$$P(u,i)=\overline{r_u}+\frac{\sum_{v\in U_N}sim(u,v)(r_{vi}-\overline{r_v})}{\sum_{v\in U_N}sim(u,v)}\ . \tag{17}$$

Step 5. Package projects are listed as scores from high to low. The recommended power selling package set P' of top N target users is selected.

7 Experiment and Analysis

7.1 Experimental Data

The experimental data came from powertochoose dataset of Texas, including 6080 power trading users, 2093 power selling package projects (detailed package published by power sellers would include the following aspects: unit price label, total electric energy, transmission and distribution expenses, electric energy expenses, prepayment, fixed charges collectable, minimum consumption, new energy proportion and package time limitation), and 1106307 star-level scores. The star-level scoring range ranged from 1 to 5. The sparseness was 96.56%.

7.2 Experimental Measurement Standards

The experiment could be mainly divided into two parts. The part one was the effective analysis of OSDA_CRWI. Effectiveness of power user subset scoring error $Err(U')$ and coverage rate of power selling package projects was used; In the part two, on the basis of finding out OFSPTU, accurate analysis of the power selling package recommendation method was conducted. Degree of Agreement (DOA), Top-K, mean absolute error (MAE) and root-mean-square error (RMSE) were used to measure good or bad recommendation effects of power selling packages.

The calculation of power trading user subset scoring error $Err(U')$ and coverage rate of power selling package projects has been introduced above, thus it wouldn't give unnecessary details here.

For the evaluation standards of the experiment two, the smaller MAE and RMSE showed the more ideal recommendation effects [20]. The computational formula of two measurement standards is shown as follows:

$$MAE=\frac{1}{|R_{test}|}\sum_{r\in R_{test}}\left|r_{real}-r_{pred}\right|, \tag{18}$$

$$RMSE=\sqrt{\frac{\sum_{r\in R_{test}}\left(r_{real}-r_{pred}\right)^2}{|R_{test}|}}\ . \tag{19}$$

In the formula, R_{test} is the test set of power selling packages; r_{real} is the real scoring of power selling package test item; and r_{pred} is the predictive scoring.

DOA is the proportion of packages with the correct sequences in all power selling packages to be calculated. $N_{U_i} = P - (L_{U_i} \cup E_{U_i})$ is defined as the package set of the training set L_{U_i} out of U_i and the test set E_{U_i} out of U_i. DOA for the user U_i can be defined as follows:

$$DOA_{U_i} = \frac{\sum_{j \in E_{U_i}, k \in N_{U_i}} Detect_{U_i}(P_j, P_k)}{|E_{U_i}| \times |N_{U_i}|} \times 100\% . \quad \textbf{(20)}$$

$Detect_{U_i}$ is the detection function and it is defined as:

$$Detect_{U_i}(P_j, P_k) = \begin{cases} 1 & if(PR_{p_j} \geq PR_{p_k}) \\ 0 & else \end{cases} . \quad \textbf{(21)}$$

Where PR_{p_j} is the ranking situation of the power selling package P_j in the recommended list. In the experiment, ultimate result of DOA is the mean of DOA in all power users. The higher DOA is, the more ideal of recommended power selling package effects will be.

Top-K means to return top K results conforming to user demands from mass data. It can be comprehended as the recommended accuracy. The higher Top-K value is, the better recommended power selling package effects will be.

7.3 Experimental Result Analysis

7.3.1 Experiment One

In order to select effectiveness of OFSPTU as verifying dataset size changes of OSDA_CRWI, some power selling package projects and star-level scoring in total dataset were selected as the test set. The remaining parts were used as the training set. In every grade with different proportions, the mean of five experiments was selected as the ultimate result. The size of OFSPTU was 50. The Genre-Based typical find algorithm (GTFA), Most Rating (MR) and Max Diversity (MD) in the literature [15] were selected for the comparative test. According to the formula (11) and formula (13), the experimental results were shown in Fig. 2 and Fig. 3.

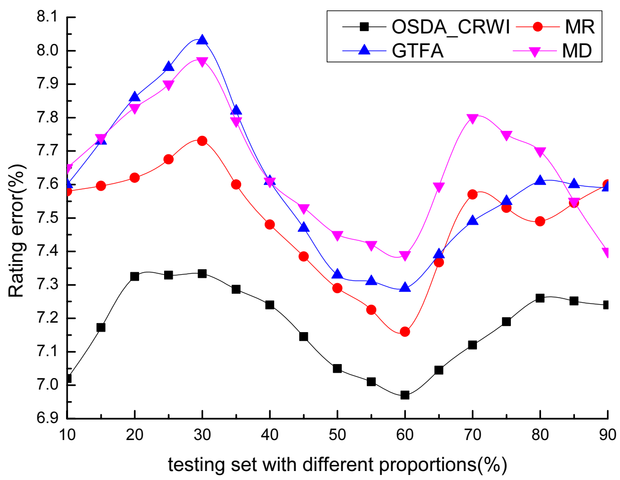

Fig. 2. Rating error of OFSPTU on testing set with different proportions

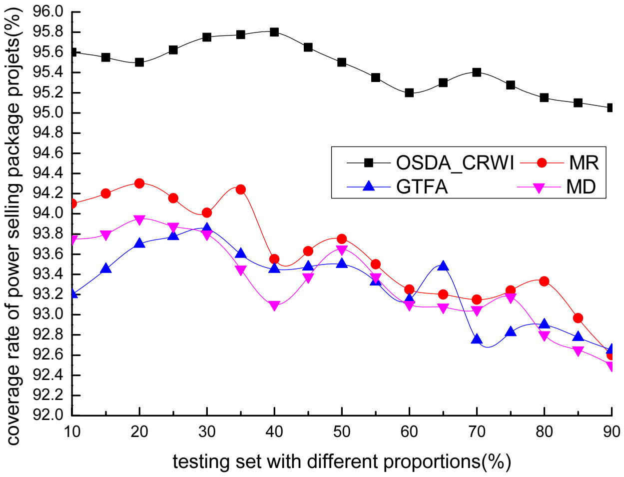

Fig. 3. Coverage rate of power selling package projects of OFSPTU on testing set with different proportions

The size of OFSPTU k could be confirmed freely. In order to verify that OSDA_CRWI could get the better result as changing optimal feature subset. 50% of power selling package projects and star-level scoring in total data set were selected as the test set. The remaining parts were considered as the training set. Similarly, mean of five experiments in OFSPTU with different size was selected as the ultimate results. The experimental results were shown in Fig. 4 and Fig. 5.

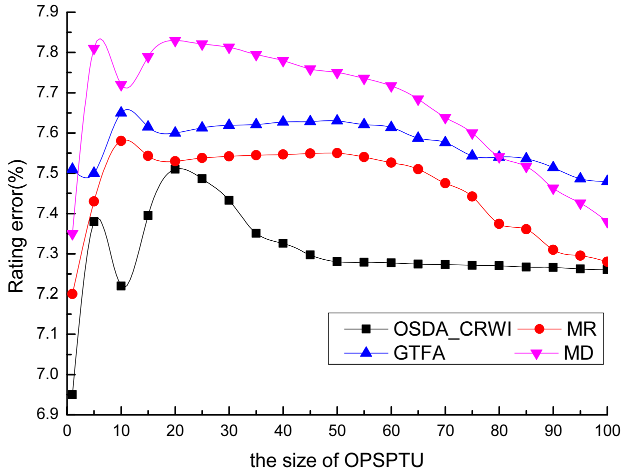

Fig. 4. Rating error of OFSPTU with different sizes

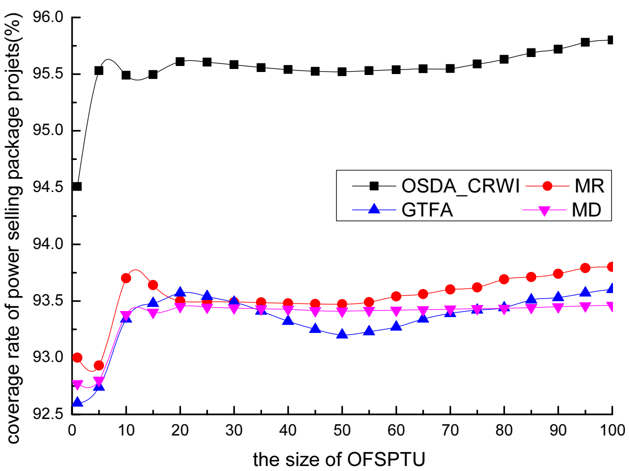

Fig. 5. Coverage rate of power selling package projects of OFSPTU with different sizes

As shown in Fig. 2 and Fig. 3, OSDA_CRWI could acquire the maximal coverage rate of power selling package projects under the circumstance of scoring error. The selected user subset could represent the interest preferences of all power trading users. GTFA algorithm didn't have the excellent performance, because the algorithm looked for the optimal feature subset according to coverage rate of power selling packages with different types. Though the feature user selected could cover every school, it didn't have representativeness. In addition, optimal subset of MR algorithm had the favorable performance, because the algorithm got the most of scores in the training set. Obviously, sets were relatively active and covered more power selling package projects. On the basis of MR algorithm, MD algorithm only considered the difference between users for users with the most of scores and the users with the smallest similarity. As a result, it couldn't effectively improve coverage rate of power selling package projects and scoring errors of power trading user subset, thus the performance was not ideal. Furthermore, when test set proportion was increased, coverage rate of OSDA_CRWI still tended to be stable, while remaining methods had a downward tendency, showing that OFSPTU selected by OSDA_CRWI had the representativeness and effectiveness.

It could be observed from Fig. 4 and Fig. 5 that as changing OFSPTU, OSDA_CRWI still obtained the good results. In addition, with the gradual increase of OFSPTU, scoring errors of several methods were increasingly closed, showing that when OFSPTU was very large, the mean user subset score has already been closing the actual average. However, the increase of power selling package projects' coverage rate was slow. The form of increasing weight reduced the possibility that unpopular power selling package projects were covered by popular projects, so OSDA_CRWI still had the great advantage in coverage rate of power selling package projects.

7.3.2 Experiment Two

In order to comprehensively verify effectiveness of PSPO algorithm, four classical recommendation methods were selected for the comparative experiment. Local user-based collaborative filtering (LUCF) algorithm in the literature [21] and local singular value decomposition (LSVD) algorithm in the literature [10] were selected. In essence, LUCF was the user-based collaborative filtering (UCF), LSVD was based on singular value decomposition (SVD). Also, the original algorithms UCF and SVD based on the package hierarchy were used for comparison. The total data set in the experiment was divided into the training set and testing set from 10% to 90% at the proportion of the total data set. The mean of five experiments on every grade was selected as the ultimate result. The size of OFSPTU was 50. At the same time, in order to comprehensively evaluate PSPO algorithm from different angles, MAE, RMSE, DOA and Top-K evaluation standards were selected. They are common evaluation methods of the recommendation algorithm.

The computed results of MAE and RMSE were shown in Table 1 and Table 2. Under most of circumstances, PSPO package scoring and predictive effects were ideal, showing the nearest neighbor of users in OFSPTU. Under the circumstance of sufficiently considering coverage rate of user subset for package recommendation item, the favorable recommendation result could be obtained. SVD, LUCF and LSVD method also indicated that user optimal feature subset acquired by OSDA_CRWI had the higher improvement effects. The typicality and representativeness were almost the best. In addition, when test set proportion was large, MAE of PSPO was basically the same with UCF, because of the large sparseness in original data, showing that OFSPTU representativeness was also determined by completeness of user data. Actually, from the proportional experimental results, recommended effects of PSPO were optimal.

Table 1. The MAE result of testing set with different proportions

testing set with different proportions /%	MAE				
	PSPO	SVD	LUCF	LSVD	UCF
10	0.8047	0.8184	0.8124	0.8148	0.8215
20	0.8018	0.8163	0.8116	0.8106	0.8285
30	0.8135	0.8262	0.823	0.8228	0.8365
40	0.8207	0.8326	0.828	0.8282	0.8463
50	0.8331	0.849	0.846	0.8423	0.8608
60	0.8498	0.8653	0.8654	0.8654	0.8726
70	0.8726	0.8823	0.8841	0.8853	0.8785
80	0.9002	0.9136	0.9171	0.9147	0.9101
90	0.9271	0.942	0.9503	0.9523	0.928

Table 2. The RMSE result of testing set with different proportions

testing set with different proportions /%	RMSE				
	SPRA	SVD	LUCF	LSVD	UCF
10	1.0204	1.0322	1.0248	1.0276	1.0519
20	1.0245	1.0334	1.0301	1.0303	1.0648
30	1.0379	1.0451	1.0411	1.0418	1.0769
40	1.0462	1.0546	1.0498	1.0504	1.0895
50	1.064	1.0769	1.0725	1.0709	1.112
60	1.0878	1.1014	1.0996	1.0998	1.1313
70	1.1238	1.1303	1.1311	1.1325	1.136
80	1.1694	1.1795	1.1828	1.18	1.17
90	1.2233	1.2336	1.2424	1.2451	1.21

The comparative results of DOA for five algorithms were shown in Table 3, so as to prove recommendation effects. It could be observed that DOA in PSPO algorithm was obviously higher than other algorithms, acquiring the better recommended effects of the power selling packages. Secondly, the recommended package results by adding business hierarchy information algorithms (LUCF, LSVD) were better than UCF and SVD, because UCF and SVD only used the common power selling package information and it would be hard to find out the neighbor user with the high reliability.

Table 3. Result of DOA comparison between five methods

Alg	UCF	SVD	LUCF	LSVD	PSPO
DOA/%	68.69	67.88	88.18	86.32	91.34

The experimental results of Top-K were shown in Fig. 6. It could be observed from the comparative experiment of the recommended method that accuracy of recommended package scoring and ranking accuracy of recommended results in PSPO were superior to other methods. The ranking results based on PSPO could rank power selling package projects favored by users in the top position, showing that PSPO could improve the recommended effects effectively.

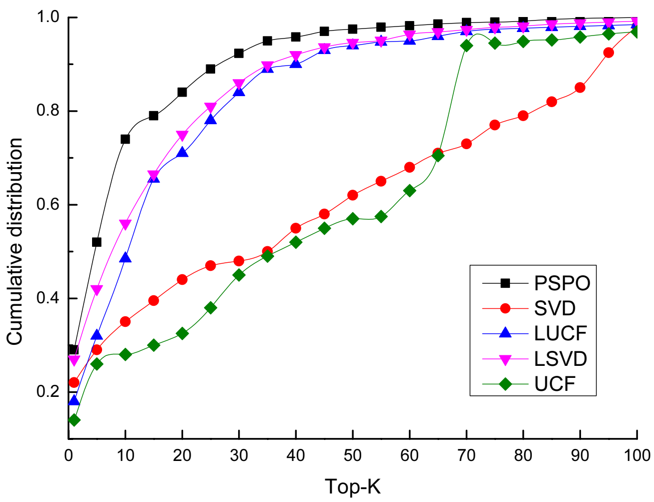

Fig. 6. Top-K recommendation of different algorithms

The running time of different recommended algorithms on the power selling package data set was shown in Fig. 7. The running time of PSPO algorithm was the shortest, because optimal feature subset of power users and package project clustering reduced the searching range in the recommended project, saving some time efficiency. The running time of SVD and LSVD was shorter than the traditional UCF and UCF-oriented LUCF, proving the effectiveness of OFSPTU in power selling package recommendation.

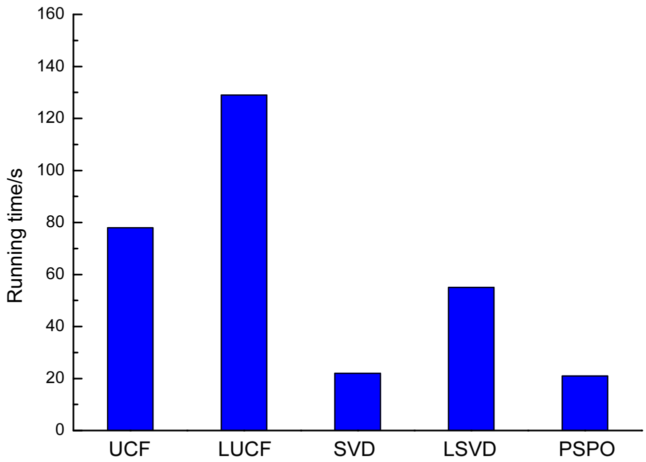

Fig. 7. Running time performance of different recommended algorithms

8 Conclusions

In the paper, OFSPTU obtained by screening out coverage rate with weighing increase items was used to stand for interest degree of all users in the package trading process. Moreover, according to the power selling package attribution in the recommended process, the similarity calculation was conducted to intensify the relationship between power selling package projects or power trading users. In the end, a recommendation method based on power selling packages of OFSPTU was proposed to recommend power selling packages to users in real time and solve the difficult user selection. The above-mentioned work could reinforce matching and application of the individual recommendation method in the power industry and enrich analysis dimension of power service products. Next, how to further reduce influences of data sparseness on optimal feature subset in the recommendation process of power selling packages will be further studied to enhance recommended accuracy.

Acknowledgements

This work is supported by he National Natural Science Foundation of China (Grant No. 51437003) and the Key Projects for Transformation of Science and Technology of Jilin Province (Grant No. 20160623004TC). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers and editors.

References

[1] X.X. Zhang, S. Xue, S. Yang, International experience and lessons in power sales side market liberalization, Automation of Electric Power Systems 40(9)(2016) 1-8.

[2] C.H. Yin, D.C. Yang, G.F. Geng, Z. Fan, Summary of energy internet projects in germany and its enlightenment to China, Power System Technology 39(11)(2015) 3040-3049.

[3] J.X. Duan, Internet innovation on the selling side, Low Carbon World 17(2016) 39-40.

[4] Y. Bai, L. Xie, Q. Xia, Q.X. Chen, H.W. Zhong, Institutional design of Chinese retail electricity market reform and related suggestions, Automation of Electric Power Systems 39(14)(2015) 1-7.

[5] Y.Z. Cheng, Analysis of the perfect Dezhou 1820 selling packages is what kind of experence. <http://www.escn.com.cn/news/show-300044.html>, 2016 (accessed 17.05.12).

[6] Y.P. Zhang, J.N. Li, H. Fu, UK retail electricity market reform and challenges, Automation of Electric Power Systems 40(11)(2016) 10-16.

[7] X.K. Wu, B. Cheng, J.-L. Chen, Collaborative filtering service recommendation based on a novel similarity computation method, IEEE Transactions on Services Computing 10(3)(2017) 352-365.

[8] W. Luan, G. Liu, C. Jiang, Partition-based collaborative tensor factorization for POI recommendation, IEEE/CAA Journal of Automatica Sinica 4(3)(2017) 437-446.

[9] X. Zhang, F.Z. Sun, C. Fang, R. Guo, Time-sensitive collaborative filtering algorithm with interest stability, International Computer Engineering and Applications 1(2017) 1-7.

[10] T. Zhang, G.J. He, Applying personalized tag group recommendation model to social photo sharing, Journal of Computers (Taiwan) 22(4)(2012) 59-67.

[11] H.Y. Wang, W.B. Yang, S.C. Wang, S.R. Li, A service recommendation method based on trustworthy community, Chinese Journal of Computers 37(2)(2014) 301-311.

[12] H. Yu, J.H. Li. Algorithm to solve the cold-start problem in new item recommendations, Journal of Software 26(6)(2015) 1395-1408.

[13] Q. Liu, E. Chen, H. Xiong, Y. Ge, Z. Li, X. Wu, A cocktail approach for travel package recommendation, IEEE Transactions on Knowledge & Data Engineering 26(2)(2014) 278-293.

[14] C. Tan, Q. Liu, E. Chen, Object-oriented travel package recommendation, Acm Transactions on Intelligent Systems & Technology 5(3)(2014) 1-26.

[15] C. Tan, Q. Liu, L. Wu, H.P. Ma, B. Long, Finding and applying typical user group in recommender systems, Pattern Recognition and Artificial Intelligence 28(5)(2015) 462-471.

[16] J.D. West, I. Wesley-Smith, C.T. Bergstrom, A recommendation system based on hierarchical clustering of an article-level citation network, IEEE Transactions on Big Data 2(2)(2016) 113-123.

[17] D. Zhang, C.H. Hsu, M. Chen, Cold-start recommendation using Bi-Clustering and fusion for Large-Scale social recommender systems, IEEE Transactions on Emerging Topics in Computing 2(2)(2014) 239-250.

[18] C. Hu, N. Luo, Q. Zhao, Fuzzy trajectory clustering technique based on fast reduced measure of semantic similarity, Journal of Computers (Taiwan) 27(2)(2016) 79-85.

[19] K.J. Kim, H. Ahn, A recommender system using GA K-means clustering in an online shopping marke, Expert Systems with Applications 34(2)(2008) 1200-1209.

[20] B.Y. Zou, C.P. Li, L. W. Tan, Social recommendations based on user trust and tensor factorization, Journal of Software 25(12)(2014) 2852-2864.

[21] H.S. Ma, Y.G. Shi, S.B. Gao, A recommendation algorithm based on collaborative filtering by feature augmentation and cascade tactics, Science Technology and Engineering 16(21)(2016) 272-277.