A novel Customer Service Recommendation Algorithm for Power Users

Yuejiao Ma*, Yongjun Liu and Ji Tao
Huaneng Lancang River Hydropower Co., Ltd, Kunming City, China

*Corresponding author email: Cocafriend@sina.com

Abstract. Aiming at the large amount of power user data and the fact that the collaborative filtering recommendation technology fails to consider the relationship between users and customer service staff, a k-means clustering and user portrait recommendation method is proposed. This method firstly uses clustering technology clustering the power users' portrait label vectors to gather similar users together, and makes recommendations based on the cluster to which the users belong. Secondly, through calculating the user portrait label vector similarity between the attributes of the service feature vector to establish the connection between the user and the customer service, and improve the traditional grading forecast method, the user and the personnel of the service of similarity index is integrated into it. Finally, the personnel of the service will be recommended from the two aspects of service quality and service fits.

1. Introduction

With the development of the recommendation system (RS), the recommendation technology used is mainly divided into three different types, namely collaborative filtering, content-based and hybrid recommendation. Among them, collaborative filtering (CF), which is one of the most successful recommendation algorithms, plays an important role in RS. CF recommendation technology collects and analyzes user's behavior preference information (such as ratings, etc.) through a large number of collections and analysis, and Use similar user or similar project information to predict other projects. According to different filtering methods, CF can be divided into memory-based CF and model-based CF [1][2].

The customer service recommendation for power users mainly has the following characteristics: (1) The amount of data information of power users is large. (2) There are intra-regional similarities and differences between regions in power user service demand. (3) There is a certain correlation between power users and customer service personnel. Although the application fields of collaborative filtering algorithms[3,4,5,6] are very extensive and the technology is very mature, there are still certain problems in applying them to power customer service. (1) The collaborative filtering algorithm is only based on the user's rating of the item to find similar users of the target user, recommending items of interest to the similar user to the target user, without considering the user and item characteristics. However, according to the actual situation in the field of electricity customer service, users have little evaluation data for customer service personnel, and it cannot be guaranteed that each user will make a corresponding evaluation after the customer service staff provides the service, nor can they guarantee that the evaluation made by each user is Objective and fair, if some users are easy to be satisfied, the customer service staff evaluation is basically 4 points, some users are more rigorous, basically the evaluation information is in the range of 1-2 points. (2) The collaborative filtering algorithm ignores
the potential correlation between users and projects\textsuperscript{[7,8,9]}. In the power customer service recommendation process, there is a correlation between users and customer service personnel in the region-region, value level-service level, because each region There is a certain difference in service indicators. The customer service staff will have a better understanding of the needs of users in similar areas, and in order to avoid the loss of high-quality users, it is necessary to recommend customer service personnel with high service levels to users with high value levels\textsuperscript{[10,11]}.

In this paper, we proposed Recommendation Model Based on the main problems mentioned above. Due to the similarity of service requirements, service indicators, similar types, and similar value levels of service users and professional literacy needs of users in the same or similar regions, user portrait technology and clustering technology can be combined. Through calculating the user portrait label vector similarity between the attributes of the service feature vector to establish the connection between the user and the customer service, and improve the traditional grading forecast method, the user and the personnel of the service of similarity index is integrated into it. Finally, the personnel of the service will be recommended from the two aspects of service quality and service fits.

2. Formatting the Text

2.1. Framework of Recommendation Algorithm Based on Clustering and User Portrait
In view of the large amount of data of power users and the fact that collaborative filtering does not consider the relationship between users and customer service personnel: a method based on k-means clustering and user portrait recommendation is proposed, referred to as framework of KUPCF. Faced with the application scenario of recommending customer service personnel for power users, there is a correlation between users and customer service personnel in regions-regions, value levels-service levels, and collaborative filtering recommendation methods ignore the potential association between users and customer service personnel. Therefore, first, cluster the power user portrait label vectors by clustering technology, cluster similar users together, and make recommendations based on the cluster to which the target user belongs; second, by calculating the similarity between the user portrait vector and the customer service attribute feature vector Degree to establish the connection between the user and the customer service, and integrate the similarity index between the user and the customer service personnel in the scoring prediction process, and finally recommend the customer service personnel from the two perspectives of service quality and service suitability. The detail framework of KUPCF is followed as Figure 1.

![Figure 1. Framework of recommendation algorithm based on clustering and user portrait](image-url)
Based on framework of Figure 1, our model includes the following steps:

1): user portrait construction: The final result will generate a power user portrait label vector, that is, \( \text{PowerUser} = <\text{e-label}, \text{f-label}, \text{s-label}> \), the vector is composed of basic user attributes (user number, user address, user Type) and value labels (including economy, load, credit, and society) that are mined through user electricity consumption and payment behavior information.

2): User portrait clustering: The purpose of this article is to recommend matching customer service personnel to users, and due to regional differences in service indicators in the power industry, user needs in similar regions are similar, and users with high value levels provide customer service Demands have similarity and other characteristics, so k-means clustering method is used to gather users with similar types, regions and value levels.

3): Similarity calculation: Similarity calculation includes two parts: (1) Calculate the similarity between users based on the user's rating matrix for customer service. (2) Calculate the similarity between the user portrait vector and the customer service attribute feature vector.

4): Scoring customer service personnel: Based on the neighborhood-based scoring prediction method, the user-customer service similarity is integrated, and the scoring is performed from two perspectives of service quality and service suitability.

4): recommendation, based on the above scoring results, recommend highly rated customer service personnel to the user.

2.2. Construction of Power User Portrait

User portrait modeling is equivalent to labeling users. Labels can be described with concise words or numbers, which is convenient for computers and humans to better understand and apply. For the modeling of power user portraits, this paper constructs power user portraits based on the idea of mathematical statistical analysis and tags users. Model power users from the three dimensions of the basic attributes of power users, namely the demographic field, the power industry field, and the business social field, and define the model as a equation 1.\( \text{PowerUser} = <\text{e-label}, \text{f-label}, \text{s-label}> \). we uses a triple to represent the power user portrait, which represents the user's basic attribute label, power industry domain label and business social domain label. e-label stands for the user's demographic static attribute label module. The module label can be directly obtained through the user's registration information and registration information.

In the labeling module of the power industry field represented by f-label, the relationship is expressed by \( \text{tag2} = f(\text{attribute2}) \). Among them, tag2 represents the user's economy, load, loyalty, and credit value level to be tapped in the current power field. In the business and social domain label module represented by s-label, the \( \text{tag3} = f(\text{attribute3}) \) relationship is also used to represent the model. Tag3 represents the social and industry value levels, and the level is also affected by the Social attributes. However, the tag 2 of the power user domain and the tag 3 of the business social domain have been introduced to include the user's economic value, load value, social value, loyalty value, credit value, and industry value. These tags are abstract tags and cannot be obtained directly from the system. They need to be obtained indirectly by studying the user's annual electricity consumption, annual electricity growth rate, number of arrears, and other electricity consumption behaviors and payment behaviors.

2.3. User Portrait Clustering Module

Although clustering technology is widely used in recommendation algorithms, but only by constructing the user's rating matrix \( R_{mxn} \), using the user's rating vector as the clustering object, this method can reduce the data dimension to a certain extent To improve the accuracy of recommendations, but the effect is not very obvious. In this paper, based on the behavior characteristics of power users and the application scenario recommended by customer service, a new method for clustering partitions using power user portraits is proposed. This method uses the constructed user portrait vector as the feature dimension of power users class. The specific process is as follows:

User portrait k-means clustering

input: n power user portrait label vectors \( \text{PowerUser} = <\text{e-label}, \text{f-label}, \text{s-label}> \), number of clusters \( k \).
output: k clustering centers.
1. randomly selects m objects from n pieces of data as clustering centers($m_1, m_2, m_k$)
2. Use the following formula to calculate the distance from each power user portrait label vector $PowerUser_i$ to the cluster center $m_k$, and divide it to the nearest cluster center

$$d(PowerUser_i,m_k) = \sqrt{\sum_{j=1}^{d} (PowerUser_{ij} - m_{kj})^2} , i = 1 \ldots n; j = i \cdots k$$

(1)

3. Recalculate the mean of the objects in each new cluster as the new cluster center, as follows:

$$m_j = \frac{1}{N_j} \sum_{i=1}^{N_j} PowerUser_{ij} \cdots j = 1,2 \cdots k$$

(2)

1. repeats step1 and step2 until the criterion function $E$ converges and returns to the cluster center ($m_1, \ldots, m_k$).

$$E = \sum_{j=1}^{k} \sum_{i=1}^{n} ||x_{ij} - m_j||^2$$

(3)

2.4. Similarity Calculation Module
The user portrait modeling method will eventually generate a user portrait vector composed of basic attribute labels, industry domain labels, and business social attribute labels, that is, $PowerUser = <e-label, f-label, s-label>$ users that can be generated according to calculations. The similarity between portrait vectors and customer service attribute vectors represents the correlation between customer service personnel. The similarity calculation between the user portrait vector and the customer service attribute feature vector is called $\text{Sim}(u, i)$ in the following text. In this paper, the commonly used vector similarity calculation method Euclidean distance and cosine similarity calculation are used for experiment.

d (x, y) represents the distance between the two vectors, $i_x$ and $i_y$ represent the i-th feature of the x and y vectors respectively.

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

(4)

When calculating the similarity between vectors using the Euclidean distance, the equation 11 is generally used. The larger, the range is between $[0,1]$.

$$\text{EucSim}(x, y) = \frac{1}{1 + d(x, y)}$$

(5)

2.5. Fusion User-customer Service Similarity Score Prediction Module
The problem of scoring prediction has always been a concern of scholars, and it is also a key indicator for detecting the quality of the recommendation system. The data set used in the scoring prediction problem is the user rating data set. The user rating record composed of data in the form of triples (u, i, r) indicates that user u rated item i as r. In this article, the user's rating record data for customer service personnel is used. The main recommendation system The score prediction method includes score prediction based on the neighborhood of the user or the project. In this paper, the score prediction method based on the neighborhood of the user is used, and the calculation is as follows:
Where pre (u, i) represents user u's rating prediction for item i, \( S (u, k) \) is the k nearest users to user u, \( N (i) \) represents all users who have rated item i, and \( R_{vi} \) is the user v The score of item i, \( Ru \) and \( R_v \) are the user u and the average score of v, Sim-score \((u, v)\) represents the calculated similarity between users.

\[
pre(u,i) = Ru + \alpha \cdot sim(u,i) + (1 - \alpha) \cdot \frac{\sum_{v \in S(u,k) \cap N(i)} \text{sim} - \text{score}(u,v) \cdot (R_{vi} - \bar{R}_v)}{\sum_{v \in S(u,k) \cap N(i)} \text{sim} - \text{score}(u,v)}
\]

(7)

\( \alpha \) represents the fusion weight coefficient, which ranges from \([0,1]\). The score prediction result \( pre(u, i) \) is no longer determined only by the user’s rating of the customer service, but also by the correlation between the user and the customer service. The higher the \( pre(u, i) \) value, the more likely it is recommended to the user.

The specific steps of the KUPCF algorithm proposed in this paper are as follows:

**Algorithm 2: KUPCF algorithm**

**Input:** power user portrait label vector, customer service information, user-customer service score information, cluster number k, neighbor number n, fusion weight

**Output:** k clusters, recommended customer service staff information

a) The K-means algorithm clusters user portraits to obtain the cluster number of the target user u.
b) Filter out the members of the cluster where the target user u is located and the evaluation information of the customer service personnel, and construct the matrix R,
c) Similarity calculation calculates the similarity between the power user u and other users in the cluster, taking n nearest neighbor users.
d) Similarity calculates the similarity between users and customer service staff.
e) Use formula 3.16 to obtain the score prediction \( pre(u, i) \).
f) Recommend highly rated customer service personnel to target users u.

3. Experimental Results and Analysis

The experimental results are shown in Figure 2. With the increase of the clustering number k in the initial stage, the SSE declines rapidly. When the k value reaches 5, it enters the elbow position, and the SSE value decreases significantly. Considering the clustering effect and clustering efficiency comprehensively, the k value of the cluster number is set to 5 for further experiments.

**Figure 2.** The graph of experimental results corresponding to cluster number k and SSE Calculation of similarity between users and customer service personnel Sim \((u, i)\) selects EucSim calculation method and changes the calculation method of similarity between users and users.
Figure 3. Comparison result of similarity selection experiment 1
Calculation of similarity between users and customer service staff $\text{Sim}(u, i)$ selects CosSim calculation method and changes the calculation method of similarity between users and users.

Figure 4. Comparison result of similarity selection experiment 2
From Figure 3 and Figure 4, respectively, the similarity between users selects CorrSim calculation method to score the prediction accuracy is generally better than the other two methods, Figure 3 and Figure 4 comparison, the similarity between users and customer service use EucSim calculation The method score prediction accuracy is better than the CosSim method.

The experimental results of different algorithms in this experiment are to select the best experimental results after many experiments. Among them, the proposed KUPCF algorithm is based on the results of the last three experiments to further experiment, the number of clusters $k$ is set to 5, based on the score matrix user The similarity module between selects the Person correlation coefficient calculation method, based on the user portrait vector-and the customer service feature vector similarity calculation module selects the European distance calculation.

Figure 5. Performance comparison between KUPCF and different recommendation algorithms
It can be seen from Figure 5 that the proposed KUPCF algorithm has improved precision, recall, and $F$ values, and the prediction scoring error has been reduced. This is mainly due to the UCF and ICF algorithms starting from the entire experimental data set, resulting in scoring The matrix is sparse, and the recommendation accuracy is poor. The Split-merge CF algorithm improves on the traditional algorithm. The clustering method is used to reduce the dimension of the score matrix, which reduces the sparseness of the matrix. All have been greatly improved, but its clustering objects are still based on the user's rating of the project, and do not take into account the relationship between the user and the customer service staff. As a result, compared with the traditional UCF and ICF, the recommendation accuracy and score prediction accuracy have been greatly improved. Compared with
the Split-merge CF, the recommendation accuracy and score prediction accuracy have also been improved, verifying the proposed KUPCF recommendation The feasibility of the method.

4. Conclusion
In this paper, based on the research of application scenarios of power customer service recommendation, considering the large amount of data of power users and the classic collaborative filtering recommendation algorithm does not consider the relationship between users and customer service personnel, a k-means clustering and user portrait is proposed Recommended method, and introduced the design framework and main modules of the method in detail. By constructing the portrait of the user, using clustering technology to process the data, and integrating the user-customer service similarity calculation in the scoring prediction process, the final score High customer service staff recommend it to users.

References
[1] Breese J S, Heckerman D, Kadie C. Empirical Analysis of Predictive Algorithms for Collaborative Filtering[J]. uncertainty in artificial intelligence, 2013.
[2] Bobadilla J, Ortega F, Hernando A. Recommender systems survey[J]. Knowledge-based systems.,2013, 46(1): 109-132.
[3] Zhang Z, Lin H, Liu K, et al. A hybrid fuzzy-based personalized recommender system for telecom products/services[J]. Information Sciences, 2013, 235(Complete): 117-129.
[4] Zhen H, Tang T, Qirong L. Research on Personalized Application Recommendation System in the OTA System Based on Mining Users' Interests[C]. International Conference on Wireless Communications, Networking & Mobile Computing. IEEE, 2008.
[5] Han H, Gao A, Ren J, et al. An O2O Service Recommendation Algorithm Based on User Context and Trust Service[C]// IEEE Trustcom/bigdata/e Spa. IEEE, 2016.
[6] Li J, Wang Y, Wu J. Application of User-based Collaborative Filtering Recommendation Technology on Logistics Platform[C]. 2013 Sixth International Conference on Business Intelligence and Financial Engineering,2013.
[7] Bagher R C, Hassanpour H, Mashayekhi H. User trends modeling for a content-based recommender system[J]. Expert Systems with Application, 2017, 87: 209-219.
[8] Subramaniyaswamy V, Logesh R. Adaptive KNN based Recommender System through Mining of User Preferences[J]. Wireless Personal Communications, 2017, 97(2): 2229-2247.
[9] Xu Y, Yin W, Wen Z, et al. An alternating direction algorithm for matrix completion with nonnegative factors[J]. Frontiers of Mathematics in China, 2012, 7(2): 365-384.
[10] Zhengyou X, Shengwu X, Ningzhong L, et al. Hot News Recommendation System from Heterogeneous Websites Based on Bayesian Model[J]. The Scientific World Journal,2014,(2014-6-26), 2014, 2014:734351.
[11] Aalipour E, Ghazisaeedi M. Recommender system introduction for requests of cancer world[J]. 2017.