Prediction of Business User Segmentation Model Based on Customer Value

Lu Siyue, Zhang Baoqun, Zhang Lu, Xu Hui, Zhang Jianxi, Ma Longfei, Wang Peiyi
State Grid Beijing Electric Power Research Institute
Beijing, China
e-mail: lusiyue2006@126.com

Abstract—**With the continuous deepening of China's power system reform, the diversification of power customer demand is becoming more and more obvious. Carrying out customer segmentation is conducive to providing better service to customers by power companies and achieving a win-win situation for both enterprises and customers. Based on the customer segments in the power industry and the proprietary features of commercial users, this paper uses the unsupervised learning algorithm (Expectation-Maximization algorithm) in big data mining technology to build a business user behavior segmentation model to solve the problem of poor operability of combining qualitative and quantitative indicators and proposed a new theory system of subdivision for business users. Aiming at the shortcomings that the EM algorithm is sensitive to the initial parameters, the EM algorithm is proposed to improve and optimize the model and enhance the accuracy of customer segmentation.**

Keywords-power industry; customer value; index system; customer segmentation; business users

I. INTRODUCTION

The power system reform has broken the monopoly of the state on the power industry and improved the limitation of customers in power consumption choice, making power supply enterprises can no longer rely on the advantages of monopoly to retain customers [1-2]. So, how to retain power customers and expand power market share is a realistic problem that every power supply enterprise must face. It requires power supply enterprises to precisely operate customers, establish a complete customer segmentation mechanism, and provide differentiated kinds of business and service types. At present, there is lack of effective customer analysis and market analysis in the domestic power industry. Because of the lack of clear market positioning, service means tend to be homogeneous [3-6].

In view of the above problems, this paper studied the characteristics of the power industry and introduced data mining technology into customer segmentation of the electric power industry combining with the current business situation, built a business customer behavior segmentation model based on data mining for the commercial users in the electro-sort. This model regards the commercial user information data as the derived variables and original variables of the EM algorithm data input, which can improve the problem of less dimensions and poor operability of the algorithm. Finally, this paper evaluated customers correctly according to the classification results of the model, recognized valuable customers and provided theoretical support for commercial customer segmentation in the power industry.

II. THE IMPORTANCE OF CUSTOMER SEGMENTATION

Customer is the core part of any business model. In order to serve customers better, enterprises need to define customer groups with different attributes and behavioral characteristics based on collecting, classifying and analyzing customer needs in all aspects, and evaluate customer value and customer risk. According to the evaluation results, customers are divided into different categories and managed. The purpose of customer segmentation is to identify the types of customers, then find the most valuable customers and implement customer retention strategies in a targeted way to improve customer satisfaction and loyalty, especially the most valuable customers [7]. Customer segmentation is not only a unified effective recognition for customers in company internal, also used to guide the customer management strategic resources disposition of enterprise. To support the personalized service and professional marketing of enterprise, customer segmentation is of special significance in today's customer-centric business economy.

With the progress of The Times, the construction of smart grid in power enterprises is increasing, which makes the power grid data show exponential growth. It satisfies the condition of customer segmentation by applying data mining technology. Therefore, combined with the features of data mining technology, based on a large number of customer data, the power customer segmentation model based on customer value is proposed. The model can identify the behavior characteristics of different customer groups, solve the shortcomings of the traditional customer segmentation model which is lack of explanation capacity, while then meet the requirements of scientific customer cognition, risk management, personalized service, and provide effective reference for decision makers

III. DISADVANTAGES OF CURRENT CUSTOMER SEGMENTATION METHODS

A. Single Dimension

Qualitative classification method is mostly adopted in customer segmentation in power industry, while quantitative classification method is rarely used. Qualitative subdivision is based on customer scale, power load grade, electro-sort, credit rating, etc. These grades have obviously been unable to meet the needs of power customers. Moreover, the

qualitative classification methods will result in rough segmentation criteria so customers' situations will be understood incomplete and demarcation boundary will be divided unclearly [8-10].

The quantitative classification methods are classified according to the level of electricity and electricity price, the amount of electricity paid, electricity properties or industries, and the flow rate of customer devices. For example, a customer segmentation architecture is established in literature 8, but it does not combine qualitative and quantitative statistical methods. The classification results are highly subjective and have a single dimension. However, the classification according to the above indicators ignores the impact of users' own characteristics on customer segmentation, and the classification results cannot be well explained.

B. Not Enough Objective

At present, the domestic customer segmentation is mainly based on the social attributes and the economic attributes of customers, rarely combined with the characteristics of the professional work itself to make customer segmentation criteria. The existing customer segmentation standards have strong dependence on expert advice, and human intervention, making the customer segmentation accuracy low and unable to adapt to the professional work. Therefore, the objective segmentation of power customers is of great significance for a better understanding of customers and enhancing enterprise marketing services and management capabilities.

C. Weak Practicability

It is found that there is almost no customer segmentation method with the combination of qualitative and quantitative classification because of the weak practicability. Mainly reflected in two aspects. First, the existing customer segmentation is out of the market so it cannot adapt to the needs of modern power enterprises marketing activities; Second, the result of subdivision is fuzzy, which cannot be well represented to the existing marketing activities of power enterprises. In addition, the traditional classification algorithm is based on non-intelligent computing and with the increasing amount of customer data, the amount of calculation increases, making it difficult for the traditional method to analyze intelligently and objectively. Literature 9, 13 and 14 put forward a customer segmentation model of the power industry based on data mining and adopted clustering algorithm such as k-means to carry out segmentation. For complex and imperfect data, the segmentation results cannot be well obtained.

IV. BUSINESS USER CUSTOMER VALUE INDEX SYSTEM CONSTRUCTION

A. Customer Segmentation Structure Model Design

A power industry business customer segmentation function structure model for commercial users was proposed based on the basis of data mining in this paper. The model consists three parts including the function, method and data. Among these three parts, data is the basic part of the whole model system. Sufficient data cleaning and processing can ensure the effective use of information mining. The correct selection of customer information field is the premise to obtain satisfactory business customer segmentation results. The method part is the most important part of the whole model. In practical application, power enterprises should flexibly select appropriate method models in combination with business to meet the corresponding needs of customer segmentation. In terms of function, it is both the cause and the result, mainly including the following modules: business customer analysis, business customer segmentation, market forecast and differentiated service.

B. Selection of Customer Behavior Indicators

In a large number of commercial customer resource information, select indicators should be selected to represent customer value. The selection of indicators should be made mainly based on business, then following other necessary design principles. In this paper, the total annual electricity consumption and electricity charges of commercial customers are used as economic indicators to avoid the interference of seasonal factors. The development index is expressed by the annual change rate of electricity consumption of commercial customers, and the ratio of used capacity and loading capacity has been used as the bearing index. Generally power industry takes the monthly net profit generated by customers as the standard to evaluate their value. High value users are not only the focus of power enterprises, but can also enjoy high-quality services and green channels. Therefore, in terms of the credit index, the change rate of payment is adopted, which is the ratio of the actual paid electricity charge of commercial customers to the electricity charge receivable. The variation index verifies whether the customer can become a stable long-term customer for power enterprises, which is of great significance for power enterprises to provide personalized services. It mainly uses the ratio of standard deviation of total commercial electricity consumption to the average value, and whether the user has the indicators of arrears or power theft.

C. Customer Value Index System

Customer value is the evaluation of the customer contribution to power supply enterprises. Customer segmentation is the analysis and evaluation of customer value in customer behavior. According to customer value, commercial users of power industry are segmented in this paper. And the value and value-added space of commercial customers to power supply enterprises can be examined in multi-dimensions. Electric power customer value is mainly composed of three parts, namely current customer value, future customer value and other customer value. Based on the experience of literature [11], the customer value index system framework is established (see Fig. 1). This framework is mainly composed of the business customer value analysis field.
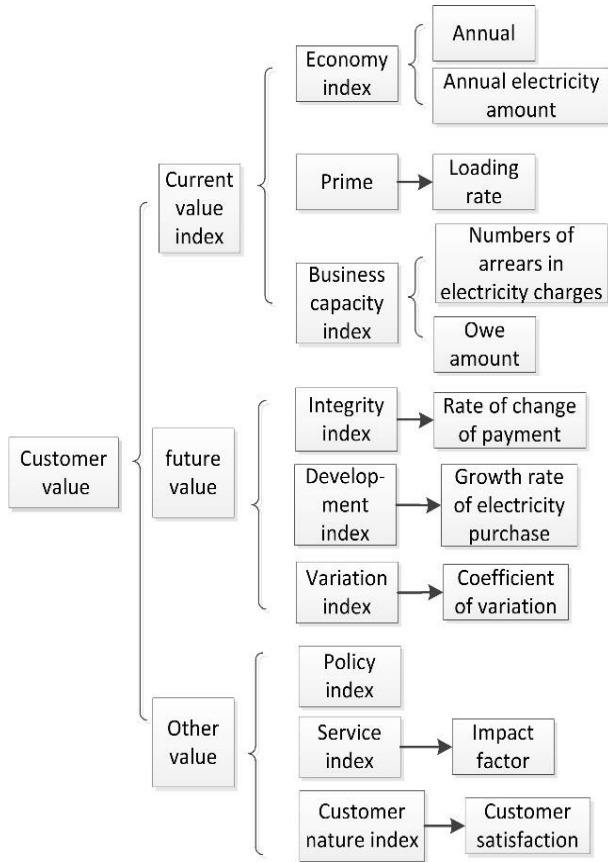
Figure 1. Customer value index system framework

D. Customer Segmentation Design Method

The general idea of business customer segmentation is mainly:

- With the combination of big data technology and business characteristics, different business customer behavior indexes are selected, and the attribute indexes of customer segmentation are continuously optimized on the basis of mining.
- Quality of customer segmentation results is assessed by the relevant indicators of customer value.
- The indicators of customer behavior are used as the basis for business customer segmentation.
- The principles of relative independence, combination of qualitative and quantitative analysis, comparability and measurability are followed [12].

E. Model Design Based on EM Algorithm

In many fields, customer segmentation usually adopts cluster analysis [13-14], logistic regression, BSC [15], EM, CRM [16-17] Artificial neural network, decision tree [18], the Bayesian [19], CLV [20], association analysis algorithm, RFM model [12] and customer value matrix. The Expectation-Maximization classification model is used to subdivide commercial customers based on the characteristics of massive data and commercial user resource information in power system.

Expectation-Maximization algorithm (EM algorithm) is a common method to solve the maximum likelihood or maximum posterior estimation problem with imperfect data. When some data is lost in the business user information, the EM algorithm uses the data of the current model to estimate the maximum posterior probability of the missing data through repeated iterative calculation, to improve the performance of the business user segmentation model.

The EM algorithm can be understood as follows:

Assume the complete data set z=(X,Y), where X and Y are observed data and missing information data respectively, As a result, z's joint density function over (x,y) on the model parameter is:

$$p(z|\theta) = p(x,y|\theta) = p(y|x,\theta)p(x|\theta), x \in X, y \in Y \quad (1)$$

The question is mainly divided into two parts:

E-step and M-step. E-step calculate the conditional expectation of the likelihood function of a complete data set, M-step perform solution maximization based on E-step result. Among them:

E-step: in data set X and current parameter set θ^{i-1}, Solve the expected value of the unobserved data set of the logarithmic likelihood function $\log P(X, Y|\theta)$.

$$Q\big(\theta, \theta(i-1)\big) = E[\log p(X, Y|\theta)|X, \theta(i-1)] \quad (2)$$

Among them, θ is the new data set after optimization.

M-step: solution maximization based on E-step result

$$\theta(i) = \arg \frac{max}{\theta} Q(\theta, \theta(i-1)) \quad (3)$$

As for the initial parameters of EM algorithm, S_k is the bounded domain of the original data feature attributes ($k=1,2\cdots d$). $W=S_1*S_2*S_3*\cdots\cdots*S_r$ form r dimensional space, divide every dimension of W by M, which means divide W into M^r cells.

The linkage between customer behavior segmentation system and customer value assessment system is a shared variable system. The key to the effective implementation of this model is to select variables that meet the needs of both customer behavior segmentation and customer value assessment. Therefore, in the framework of customer value index system, the commercial users' fields needed to be analyzed are imported into the analysis software: total electricity quantity, total electricity charge, amount of arrears, power supply unit, number of arrears, coefficient of variation, customer satisfaction score, loading capacity, used capacity and derivative variables. Among them, the fields with representative business user attributes are also taken as part of the original variables, such as the impact factors in the policy impact indicators and the industry of the customer. After cleaning and deleting abnormal and invalid data, and then eliminating dimensional influence through standardized data process steps, the final data set is the set required for final analysis.

The improved unsupervised learning EM algorithm was used to conduct consumer segment with the sorted data. The EM analysis framework was established (see Fig. 2). According to the results of visual ranking of customer value,

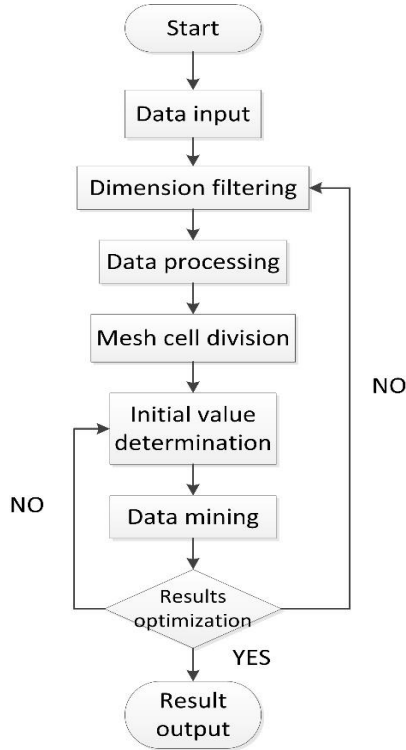business users can be divided into super high value, high value, ordinary value and low value customers.

Figure 2. EM analysis framework

F. Customer Segmentation Service Suggestions

Regarding the super high valuable customer, they are the best resources for power enterprises, and continuous growth of profits can be assured if the cooperation partnership can be made with them. In order to win more customer loyalty, "one-to-one service" should be implemented, the dynamic needs of customers should be timely grasped, and best service should be provided for such users. In the high value customer level, this kind of user determines the success or failure of power enterprises in the fierce competition, and is also the customer resources that power enterprises compete with each other. In addition, such users have high value but low development potential. In general, customer retention and frequent marketing strategies are adopted to maintain the relationship with such customers. As far as ordinary value customer is concerned, it is contrary to the characteristics of high value customer. It is a kind of customer with low value but high potential. Such customers need a lot of investment from power companies to transform from ordinary value customers to high value customers. Therefore, power enterprises need to adopt the strategy of customer training and customer intimacy. Low value customers are the lowest class of customers. Such users are not attractive to power enterprises. In order to ensure the overall business of power enterprises, the strategy of contract cancellation and appropriate transformation is generally adopted.

V. CONCLUSION

Based on the big data mining technology, the commercial user value assessment system in the power industry was discussed in this paper, and solved the problem of poor operability of the combination of qualitative and quantitative indicators. Throughout the whole paper, the following conclusions are drawn:

1. Business users need precision marketing. According to the research findings, firstly, the power industry has a complete database system, which can store a large amount of commercial user information and meet the hardware conditions for customer segmentation of data mining. Second, under the electric power reform, it is urgent to formulate refined marketing strategies.

2. The initial parameters in business user segmentation are uncertain based on EM algorithm. Aiming at this problem, a mesh EM clustering algorithm is proposed to improve the accuracy of classification results, which has certain practical value.

3. A business customer value index model based on data mining was built. In this model, existing variables are used as the basis and derivative variables are used as the link to carry out customer segmentation, so as to make up for the lack of customer value segmentation of commercial users in the power industry and provide a theoretical basis for the segmentation of commercial users and marketing work.

4. Build a customer value evaluation system based on data mining. This system evaluates customer value comprehensively through current value, future value and other values. It is an important supplement to the business customer segmentation system and can solve the problem of poor interpretation of customer behavior.

REFERENCES

[1] ZHOU Ying, LV Wei, JING Miao. Construction of large customer screening and segmentation model for power enterprises and empirical[J]. Industrial engineering and management, 2008, 23(6): 103-107.

[2] YU Hui.A Review of Customer Segmentation Methods [J] .Business Management, 2012 (2).

[3] CHENG Ci. Application of clustering analysis algorithm in power marketing decision support system[D]. Liaoning: Liaoning Technical University, 2008.

[4] LIU Xiaoxiao, GUO Xinze, TIAN Jianwei. Research on Customer Segmentation and Value Evaluation Based on Electricity Customer Behavior - Research on the Construction and Application of Customer Segmentation Model in Telecom Enterprise [J] .Electronic World, 2011, 6 (30): 245-248.

[5] Cheng Zhaoxing. Exploration of Customer Segmentation Modeling for Power Industry Based on Data Mining [D] .Chongqing University, 2007.

[6] XU Meng Festival. Power supply company first-line team bearing capacity analysis [J]. Statistics and Decision 2011, 18: 183-185.

[7] ZENG Zhongjie. Application Research of Power Customer Experience Management [D]. Huazhong University of Science and Technology, 2013.

[8] Shan Lianqiu. Analysis of Customer Segmentation of Power Supply Enterprises in the Big Data Era - A Case Study of Dongguan Power Supply Bureau [J]. Economic Management, 2016,34-35.

[9] SONG Cai-hua, LAN Yuan-juan, et al. Analysis of customer segmentation model in power industry based on data mining[J]. Model electronic technology, 2014, 37(14): 21-27.

[10] LAI Yuyang.Customer Service Strategy Based on Customer Value in Power Enterprises [J] .Research & Design, 2016,32 (14): 51-53.

[11] LI Weicheng. Electric vehicle management system design and implementation [D]. Shandong University, 2013.

[12] ZHANG Yan. Design and implementation of power customer credit risk assessment management system [D]. Beijing: North China Electric Power University, 2013.

[13] DAI Junliang. Power market application based on CR[D]. Beijing: North China Electric Power University, 2008.

[14] WANG Lei.Research on customer segmentation model of power industry based on data mining [D]. Shanghai: Shanghai Jiao Tong University, 2007.

[15] Wang Yi. Study on the Performance Evaluation of ERP Implemented by Electric Power Company [D] .Northwest University, 2013.

[16] Liang Suidong. Yulin Power Supply Company Customer Relationship Management [D]. Lanzhou University of Technology, 2013.

[17] Fang Jing. CRM-oriented data mining in power marketing applications [D]. Tianjin University, 2004.

[18] SONG Caihua, LAN Yuanjuan, FAN Ting, LI Bingtao. Application of comprehensive value evaluation method in customer segmentation of power enterprises[J]. Electronic Design Engineering, 2012014,12(22): 111-116.

[19] Zhou Hui. Electricity customer credit analysis and arrears early warning under market conditions [D]. Beijing Jiaotong University, 2010.

[20] Tang Mingrun. Distribution Marketing Research Based on CLV and Customer Satisfaction [J] .Journal of Chongqing University, 2013.