Statistical Analysis and Countermeasures of Major Power Customer Loss

Ying Zhou¹, Feixiang Gong¹, Qing Li¹, Shupeng Li², Xianxu Huo² and Dezhi Li³

¹China electric power research institute, Beijing
²State grid Tianjin electric power company, Tianjin
³China electric power research institute, Beijing

Abstract. This paper defines the loss of major power customer from the perspective of customer relationship management with a combination of qualitative analysis and quantitative analysis. The life cycle theory and risk management theory are used to analyze and discuss the factors of the loss of major power customers. The analysis model of major power customer loss based on evolutionary game is constructed, and empirical research is carried out by using the survey data. From the perspective of quantitative analysis, the method of variable selection and variable optimization selection for power customer churn early warning model is studied. Based on the comparative analysis of customer churn early warning model, a logistic regression based customer churn warning model is constructed. The effectiveness of the model is verified numerically with the survey data. Finally, the coping strategies such as differentiated service, organizational structure adaptability adjustment strategy and information system upgrade service are proposed.

Keywords: Major power customer; customer loss; early warning; coping strategies.

1. Introduction

With the gradual deepening of electric power reform, the market pattern of long-term monopoly of power grid enterprises is gradually being broken. The introduction of social capital and the formation of a new electricity market gradually have a greater impact on power grid companies, which has brought about significant changes in the business content, service objects and external situation of power grid operation. On the one hand, after the opening of the power supply side, the power grid company directly faces the pressure of losing part of the electricity sales market, resulting in a decline in market coverage and occupancy of the power grid company, and a decrease in sales profit. Especially when the power generation companies directly increase their electricity sales business or engage in the distribution of electricity business by integrating with social capital, the larger market share of high-quality market demand will be occupied, which will have an impact on the market pattern of power grid companies. In addition, the liberalization of distribution business will attract a large amount of capital investment, and even promote the establishment of an independent distribution power integration enterprise, which will also bring considerable competitive pressure to the operation of power grid companies. On the other hand, the establishment of the electricity sales market means that the seller side has evolved from the seller oriented to the buyer oriented, and the selectivity of power customers to power suppliers has been enhanced. In order to maintain the customer resources, especially the high-quality customer resources, grid companies need to balance the market customer resources with the market share of electricity sales and whether electricity bills can be recycled[1].
2. Statistical Analysis of Power Customer Loss

In order to make the statistics of the loss of major power customer representative, we first collected the data of social electricity consumption (including industry, urban and rural residential electricity), average growth rate of three-year electricity sales (%), power supply service density (power supply population/power supply area), regional GDP, regional per capita GDP, regional fixed asset investment and total retail sales of consumer goods in the whole society in 130 county-level power supplies under the jurisdiction of Sichuan Electric Power Grid Company in 2017 according to the dimensions of scale, growth, power supply service density (power supply service population/power supply service area) and regional economic development level, taking the county-level power supply company as the unit. Then the cluster analysis method and the software of SPSS 19.0 were used to analyze these data for electricity market structure classification.

According to the results of clustering analysis, the above mentioned 130 county-level companies can be divided into five types of electricity market, that is, the high-quality electricity market represented by Chengdu Qingbaijiang area with large scale, rapid growth and high power supply service density; the electricity market represented by Guanghan City with relatively high level of regional economic development and large scale of electricity consumption but with high volatility; the electricity market represented by Ganzi Tibetan Autonomous County with a small scale of electricity, a low density of power supply services, and a low level of regional economic development but rapid growth in electricity consumption; the electricity market represented by Weiyuan City, a traditional industrialized city, with a large scale of electricity consumption, large fluctuations in power consumption, and mainly a downward trend, and a general density of electricity service; and the electricity market with low economic efficiency represented by Yingjing county, which has a small scale of electricity consumption, slow growth in electricity consumption, low level of regional economic development, and low density of electricity service. The results are shown in Table 1 and Table 2.

Table 1. Five types of electricity market based on cluster analysis.

| Types of Electricity Market                                                                 | Typical representative                  |
|---------------------------------------------------------------------------------------------|----------------------------------------|
| High-quality electricity market with large scale, fast growth and high density of power supply services | Qingbaijiang Power Supply Company       |
| Electricity market with relatively high level of regional economic development, acceptable scale of electricity consumption but high volatility | Guanghan Power Supply Company          |
| Weiyuan, a traditional industrialized city, has a large scale of electricity consumption, a large fluctuation in electricity consumption growth and a downward trend, and a relatively general electricity market with a relatively high density of electricity service. | Weiyuan County Power Supply Company    |
| Most electricity markets with small scale, slow growth of electricity consumption, low level of regional economic development and low density of electricity service have low economic benefits. | Xingjing County Power Supply Company   |
| The power consumption scale is very small, the power supply service density is very small, and the regional economic development level is low, but the power consumption growth fast electricity market | Ganzi County Power Supply Company      |
Table 2. Basic characteristics of different electricity markets.

| Representative of Typical Electricity Market Types | Total electricity consumption of the whole society (10,000 KW) | Density of power supply services (10,000 people/km²)% | Average growth rate of electricity consumption in recent three years (%) | Per capital GDP (10,000 yuan per person) |
|---------------------------------------------------|---------------------------------------------------------------|-----------------------------------------------------|-----------------------------------------------------------------------|-----------------------------------------|
| Qingbaijiang Power Supply Company Class A          | 299320.38                                                     | 1.1693                                              | 12.65                                                                 | 91661                                   |
| Guanghan Power Supply Company Class B              | 170232.93                                                     | 0.1115                                              | -2.16                                                                 | 59877                                   |
| Weiyuan County Power Supply Company Class C        | 253602.11                                                     | 0.0557                                              | 1.17                                                                  | 53409                                   |
| Xingjing County Power Supply Company Class D       | 90664.56                                                      | 0.0035                                              | -7.99                                                                 | 43719                                   |
| Ganzi County Power Supply Company Class E          | 3836.25                                                       | 0.0008                                              | 33.50                                                                 | 12338                                   |

The conditions of the above five typical electricity market are highly correlated with the regional economic development level of Sichuan Province and the regional leading industry development policy. Taking Qingbaijiang area as an example, it is the undertaking area for local governments to promote key leading industries (advanced manufacturing industry, big data industry park), showing a good dual growth momentum of electricity consumption and quality. The main industries in Weiyuan County are metallurgical manufacturing and steel production bases. Influenced by the supply-side reform of steel industry, market over-saturation and industrial economic transformation in Sichuan Province in recent years, the scale of electricity consumption is large, but the fluctuation of power growth is also large. Due to the comprehensive influence of relevant industrial policies and market supply and demand situation, the production and operation of the power market around the core enterprise (a steel group) and its supporting suppliers at the upper and lower levels have been significantly impacted, showing that a large number of related power customers appear to be closing the account[3].

At the same time, the industrial agglomeration with Chengdu as the core also shows a significant siphon effect on the surrounding regional economic elements. The "4-hour" economic circle in the developed regions supported by highway network and high-speed railway rapidly covers the radius of the core industrial economic region of 500 kilometers, which accelerates the transfer of supporting industries in the surrounding regions to the core region. As a result, the industrial economic factors of the geographical area of 300 kilometers away from Chengdu, such as Yingjing county, are rapidly lost. See Table 3.
Table 3. Statistics of sales in five types of electricity market from 2007 to 2017 (excluding transitional customers).

| Industry Categories of Sales Customers | Qingbaijiang | Guanghan County | Weiyuan County | Yingjing County | Ganzi County |
|----------------------------------------|--------------|-----------------|----------------|-----------------|--------------|
| A. Agriculture, forestry, animal husbandry and fishery | 40 | 0 | 0 | 36 | 2 |
| B. Industry | 331 | 5 | 1 | 170 | 150 | 5 |
| B1 Mining industry | 86 | 1 | 36 | 49 | 0 |
| B2 Manufacturing industry | 196 | 5 | 0 | 90 | 101 | 0 |
| B3 Production and Supply of Electricity, Gas and Water | 50 | 0 | 0 | 44 | 1 | 5 |
| C. Construction business | 119 | 0 | 2 | 49 | 49 | 19 |
| D. Transportation, Warehousing and Postal Service | 8 | 0 | 0 | 8 | 0 | 0 |
| E. Information transmission, computer services and software | 5 | 0 | 0 | 5 | 0 | 0 |
| F. Commerce, accommodation and catering | 63 | 0 | 0 | 36 | 0 | 27 |
| G. Finance, Real Estate, Business and Resident Services | 13 | 2 | 0 | 11 | 0 | 0 |
| H. Public utilities and management organizations | 31 | 0 | 0 | 26 | 0 | 5 |
| Total | 782 | 7 | 4 | 511 | 202 | 58 |

Price factors, product quality factors and service factors are the core elements of customer churn. The current top-level design of the electricity trading market and the pilot reform of incremental distribution network is still in the process of exploration, adjustment and improvement, and the electricity pricing mechanism generated by the trading game between the market supplier and the demand side has not yet been formed. The government-led administrative pricing features are remarkable, and it is more difficult to obtain complete and real electricity market transaction data of the power producers and purchaser in the electricity trading market in a short time. In this case, it is difficult to quantitatively analyze the external risk factors of power customer churn. On the other hand, external factors, such as the adjustment of national industrial policy, the future Sino-US trade war and the incremental distribution network reform, which are reflected by major power customers, are mainly based on the changes in electricity consumption. In addition, together with the internal service factors that affect the loss of major power customers, such as the expansion of packaging quality, marketing service level, and other energy substitutions, it constitutes the main risk factor for the loss of major power customers[4].
3. Construction and Empirical Results of Early Warning Model for Power Customer Loss

In the experiment, 60% of the sample set was used to train the model and 40% was used to test the model. The total number of customers selected from power grid companies was 403, of which 311 were normal customers, 92 were lost customers and the proportion of customers lost was 22.7%. Using the process of variable selection in section 3.1, 12 variables in the original data are processed to obtain 6 new variables, so that the standardized scoring coefficient matrix can be obtained. Since the six common factors obtained by this calculation are mutually orthogonal, there is no multiple collinearity relationship.

Table 4 shows the confusion matrix in the initial stage of logistic regression analysis. It can be seen that 311 customers actually did not lose and the model predicted correctly, with a correct rate of 100%; 92 customers actually lost and the model predicted incorrectly, with a correct rate of 0%, and the overall prediction accuracy of the model was 77.3%.

| Actual measurement | Forecast | Loss or not | Correct percentage |
|--------------------|----------|-------------|---------------------|
| Loss or not        | 0.00     | 311         | 0.00 100.00         |
| Overall            | 1.00     | 92          | 0.00 0.00           |
|                    |          |             | Overall percentage 77.30 |

From the variables shown in Table 5 which are not included in the equation, we can see that the minimum observed value of score test statistics is 4.667, and the significance level of all variables is less than 0.05. Therefore, it is of statistical significance for the whole model fitting optimization to incorporate the above six variables into the model.

| Step   | -2Log likelihood | Cox and Snell $R^2$ | Nagelkerke $R^2$ |
|--------|------------------|---------------------|------------------|
| 1      | 8.535$^2$        | 0.429               | 0.700            |

As can be seen from the classification table in Table 7, 0 indicates no loss and 1 indicates customer loss. The total number of customers selected by the model is 403, of which 92 are actually lost and 311 are not lost. Among the 92 lost people, 64 were correctly identified by the model and 26 were incorrectly identified, with a correct rate of 19.1%. The total prediction accuracy of the model is 86.4%.
Table 7. Classification Table.

| Actual measurement | Forecast | Loss or not | Correct percentage |
|---------------------|----------|-------------|--------------------|
|                     |          | 0.00 | 1.00 |
| step2               | Loss or not | 0.00 | 311 | 28 | 90.80 |
|                     | 1.00 | 92 | 64 | 71.10 |
| Overall percentage  |           |           | 86.40 |

According to the variables in the equation shown in Table 8, we can know that the significant level of each variable is below 0.05. Based on these results, we can get the binary regression equation model:

$$\ln \left( \frac{p}{1-p} \right) = -3.719 - 1.781F_1 + 0.939F_2 + 1.715F_3 + 1.953F_4 + 0.090F_5 - 1.125F_6$$ (1)

Table 8. Variable parameters of binary regression equation.

| Variable | Coefficient B | Standard error | Wald | Freedom | Saliency | Exp(B) | 95% confidence interval |
|----------|---------------|----------------|------|---------|----------|--------|------------------------|
|          |               |                |      |         |          |        | lower limit | Upper limit |
| F1       | -1.781        | 0.707          | 6.343 | 1       | 0.012    | 0.168  | 0.042 | 0.674       |
| F2       | 0.939         | 0.386          | 5.918 | 1       | 0.015    | 0.391  | 0.183 | 0.833       |
| F3       | 1.715         | 0.779          | 4.850 | 1       | 0.028    | 0.180  | 0.093 | 0.828       |
| F4       | 1.953         | 0.322          | 36.692 | 1 | 0.000    | 7.048  | 3.747 | 13.259      |
| F5       | 0.909         | 0.412          | 4.880 | 1       | 0.027    | 0.403  | 0.180 | 0.903       |
| F6       | -1.125        | 0.231          | 23.783 | 1 | 0.000    | 0.0325 | 0.206 | 0.510       |
| constant | -3.713        | 1.671          | 4.939 | 1       | 0.026    | 0.024  |        |             |

Through the binary regression loss warning model of grid customers, the confusion matrix of grid company customers can be obtained. The results are shown in Table 9.

Table 9. Empirical Assessment Matrix.

| Customer status | Forecast loss | Predicting non-loss | Total |
|-----------------|---------------|---------------------|-------|
| Actual loss     | TP            | FN                  | TP+FN |
| Actual non-loss | FP            | TN                  | FP+TN |
| Total           | TP+FP         | FN+TN               | PF+TN+FN+TN |

After forecasting by customer churn model, customer samples can be divided into potential churn group and non-potential churn group. Among them, TP is the number of samples that the model correctly predicts the lost customers as lost customers; TP model correctly predicts lost customers as the number of lost customers; FP is the number of samples that the model predicts non-lost customers as lost customers; FN is the number of samples that the model predicts lost customers as non-lost customers; TN is the number of samples that the model can accurately predict the non-loss customers as non-loss customers.
Table 10. Empirical Assessment Data.

| Customer status    | Forecast loss | Predicting non-loss |
|--------------------|---------------|---------------------|
| Actual loss        | 9             | 1                   |
| Actual non-loss    | 5             | 164                 |

From the confusion matrix obtained above, it can be seen that the actual number of customers not lost is 169. Among the 10 power grid customers who actually lost, the model correctly identified 9 customers and misidentified 1 customer, with a correct rate of 90%. The overall prediction accuracy of the model is 96.3%.

In the prediction of customer churn model, hit rate and coverage rate are two important evaluation indexes. Hit rate = correctly analyzing the number of lost customer samples hit / predicting the total number of lost customer samples; Coverage rate = correctly analyzing the number of lost customer samples hit / the total number of lost samples; Accuracy rate = correctly predicting the number of lost and non-lost samples / the total number of all samples. Therefore, the hit rate and coverage of the customer churn model are obtained as shown in Table 11.

Hit Rate = \( \frac{TP}{TP+FP} \)  
Coverage = \( \frac{TP}{TP+FN} \)  
Accuracy = \( \frac{TP+TN}{TP+FN+FP+TN} \)

Table 11. Empirical Assessment Results.

| Empirical results | Accuracy rate | Hit rate | Coverage rate |
|-------------------|---------------|----------|---------------|
|                   | 96.3%         | 64.3%    | 90%           |

Through the evaluation of the experimental results, it can be seen that the built power grid company customer churn early warning model is significant in terms of model accuracy, hit rate, coverage and promotion. As the new power reform is still in its initial stage, the proportion of customer churn in power grid companies is relatively small at this stage among a large number of customers. Therefore, besides the above specific numerical analysis, the results of empirical analysis can also be introduced into the ROC curve of power grid companies to reflect the practicability of the empirical results of customer loss in power grid enterprises.

ROC curve is based on a series of different binary classification methods. The true positive rate (i.e. sensitivity: the probability of judging the actual value as the true value) is ordinate, and the false positive rate is 1-specificity: that is, the probability of judging the actual value as the false value is drawn for abscissa[5]. The closer the ROC curve is to the upper left corner, the better the accuracy of the model is.

4. Coping Strategies of Power Grid Company under the Background of New Electricity Reform

On the one hand, the reward or punishment obtained by the reciprocity and cooperation strategy of major power customers should be designed as explicit monetary reward incentives or punitive constraints to guide the behavior of major power customers; on the other hand, by strengthening the grid company's service brand image, increasing the content of power service, and enhancing the customer service perception experience, such as hidden non-monetary incentives, the satisfaction of power customers is enhanced, and loyalty to enterprises is fostered. Finally, the two incentive mechanisms support and complement each other, and guide large power customers to take more reciprocal and cooperative strategies in participating in power market transactions dominated by power grid enterprises, so as to form a long-term stable cooperative relationship between large customers and power grid enterprises. At present, power grid enterprises should take precautions and take the following actions: launching a special study on the loss of large power customers in different electricity markets; actively participating in the pilot project of incremental distribution reform; formulating a systematic strategy for promoting the service of power major customers[6].
5. Summary

Under the new situation, how to integrate the commercial attributes of power services and products with the operation mechanism of the new power market economy, constantly adapt to new external environmental changes, retain the old customers by innovating and utilizing various marketing methods, strive for new customers, actively expand the power market resources under the new situation, and form sustainable development power are important practical problems that power grid enterprises need to solve in the future [7]. At present, in the process of building a "customer-centered" modern service system, power grid enterprises should strengthen demand-side research and innovation management, and enhance the ability to control the dynamic changes of power marketing market in real time. Among them, it is particularly necessary to actively study the strategies to cope with the loss of high-quality power customers and the double pressures brought by competition with power selling companies due to the establishment of new electricity market. Furthermore, electricity price level and electricity service experience are the main considerations for power customers to choose power supply service objects under the conditions of electricity market. Power grid enterprises need to make use of abundant and powerful customer basic information data to conduct targeted research on customer churn and predict power customer churn, so as to achieve the goal [8]. It provides decision-making basis for formulating corresponding service strategies and striving for incremental market share.

Acknowledgment

This work is supported by the Science and technology projects from State Grid Corporation: Research and Application of Key Technologies on Regenerative Electric Heating Optimal Configuration and Interactive Control (SGTJK00DWJS1900101).

References

[1] Yang C, Li X, Ren Y, et al. Statistical Analysis and Countermeasures of Gas Explosion Accident in Coal Mines [J]. Procedia Engineering, 2014, 84(84):166-171.

[2] Park D, Perros H G, Yamashita H. Approximate analysis of discrete-time tandem queueing networks with bursty and correlated input traffic and customer loss [J]. Operations Research Letters, 1994, 15(2):95–104.

[3] Olivier J. T. Briët, Yukich J O, Pfeiffer C, et al. The effect of small solar powered ‘B̄k̄s̄’ net fans on mosquito net use: results from a randomized controlled cross-over trial in southern Ghana [J]. Malaria Journal, 2017, 16(1):12.

[4] Yangjing. Research on the Management Strategy of Power Supply Enterprises Against the Loss of Large Users [J]. Manager, 2012(10): 191-191. yangjing. Research on the Management Strategy of Power Supply Enterprises Against the Loss of Large Users [J]. Management, 2012(10): 191-191.

[5] Lishuguo, Lichujun. A new method of data mining of user loss process of power supply service channel [J]. China Science and Technology Information, 2018, No. 575(2): 90-92.

[6] Hanxiangbin. The status quo and countermeasures of the loss of basic electricity charges for large industrial customers in Wuhan [J]. Power Demand Side Management, 2006, 8(6): 51-51.

[7] Jiangyun, Li Dong, Liuzhaozhang. Marketing risk control measures under the situation of power system reform [J]. Modern Marketing (later issue), 2017(2): 45-45.

[8] Pisarenko V F, Rodkin M V. Statistical Analysis of Natural Disasters and Related Losses [J]. Springerbriefs in Earth Sciences, 2014.