Model Construction of Early Warning for Frequently Outage Complaint Based on Data Mining

Jun Fu^{1,2}, Xin Xu^{1,2}, Zhijie Sun^{1,2}, Li Wang^{1,2}, Dongmei Gong^{1,2}, Lingyu Zhang^{1,2}

¹Electric Power Research Institute of State Grid JIBEI Electric Power Company, Beijing

²North China Electric Power Research Institute CO.,LTD., Beijing 100045

Abstract. At present, frequent outages have become the major source of power customer complaints and seriously affect improvement of customer service satisfaction. The current control of frequent outages complaints has been in a passive state of compensation, which can only get half the result with twice the effort and has caused adverse perception of customers. In response to this problem, this article takes the initiative to prevent as a starting point, through studying rules of complaints business for the frequent power outage, constructs the early warning model of the frequent outages complaints, which takes statistics of outages as the data mining object and uses Chinese word segmentation matching algorithm as data mining technology and displays through network map technology, and eventually realizes a frequent power outage scientific and accurate complaints warning. It provides an accurate reference for properly arrange of troubleshooting and maintenance of the power supply enterprises and line reconstruction plans, offers technical support for the forward service gateway and lays a solid foundation for effectively reducing the volume of complaints.

1 Introduction

Frequent outages complaints have been the major source of complaints from power customers. According to the complaint statistics provided by State Grid JIBEI Electric Power Company, frequent power outages accounts for about 40% of the total number of complaints, and even accounts for 50% of the total number of complaints in a few years.

At present, the main basis for determining whether a frequent power outage complaint is is to manually check the customer appeal. Such method, which can only passively analyze and process after customer appeal emerged, can only get half the result with twice the effort and is not conducive to improve of service quality.

How to correlate fault repair information, power outage information and frequent power outage complaints that have been generated to form an early warning mechanism has always been the focus of power supply enterprises. According to the business rules of frequent power outage complaints, the address information is extracted from the generated fault repair information and outage information to count the number of power outages of a power outage unit, and the early warning information is displayed with the cloud map so as to achieve the purpose of early warning for frequent outages.

2 Early warning model for frequent power outage based on data mining

2.1 Principle Of Early Warning Model For Frequent Outages

According to the definition of complaint Type in Measures for the Management of 95598 Customer Service Business of State Grid Corporation of China, frequent outages refer to the problems of frequent power outages or long-term unimproved improvements and incomplete handling. Specific details of the operation is that the customer reacts two months of outage three times or more (including plans, temporary, power failures); The customers, who are strongly dissatisfied with the frequent power outages and require definitely complaints, should be classified as frequent power outages. Therefore, this early warning model takes outage, which lasts for two months(calculated according to 60 days in two months)three times in a single statistical area, as the critical point for early warning of frequent blackouts.

It can be seen from the above that if the start time and the end time of mining the fault repair data and the power outage information data in each statistical area are regarded as the time points, the ending time is assumed to be $t_0 + 60$, the starting time is assumed to be t_0 . The start time and recovery time of fault repair data and outage information data are T_s and T_f . Take

* Corresponding author: 676269930@qq.com

$i \neq k, i \leq n; n, i, k \in N$ as the data mining interval, when there is a minimum value, the condition of a single outage event is established as $f(k, i) = (Ts_i > Tf_k) = 1$, that outage event does not establish is marked as: $f(k, i) = (Ts_i < Tf_k) = 0$, take this number outages in data mining cycle as S . The results of the project will include: business, understanding appeals which urgently need to be solved in customer service, so $S = \sum_{k=1}^n f(k, i)_k$, thereinto $i \neq k, i \leq n; n, i, k \in N$. The above formula can rule out the crossed blackout events at the same time window, and calculate the number of power outages in each data mining cycle, and then extract the number of power outages used in frequent power outages complaints early warning, and achieve early warning map display.

The main features of early warning for frequent blackouts complaints include: early warning data management, early warning data mining, early warning rules management. Thereinto:

(1) 95598 work order data management can import existing work order data, update regularly, save, inquire and other operations.

(2) By defining data mining rules, early warning data mining control algorithm execution process, realize statistics of the number of power outages, and finally output the work order data can be used for early warning manifestation.

(3) Pre-warning rule management is to define the pre-warning level and early-warning content.

Fig.1. process of early warning of frequent power outage

2.2 Process of early warning data mining

Early warning data mining mainly includes three parts.

2.2.1 Create an early warning database

Through the interface with the 95598 work order database, the trouble ticket data and the power outage information data are imported and preprocessed to generate the original data source, which forms multi-dimensional early warning data for word segmentation, matching and analysis and mining. The so-called early warning data processing is mainly for fault work order data and power outage information data, adjusts and intercepts in accordance with the provisions of the

format field, which is conducive to facilitating subsequent data mining.

Create early warning database mainly from two aspects: First, define early warning data field: there are a number of different fields in two kinds of work order information data of fault work order data and blackout information data, but its meaning is basically the same. Therefore, these fields can be categorized as the same field. For example: the processing time and the repair time in the trouble work data has the same business meaning as the start power failure time and the restoration power supply time in the power failure information data. Therefore, during the data preprocessing, the acceptance time and the power-off start time, repair time and power recovery time are grouped into the same fields. Second, limit the scope of data mining: According to the existing early warning model, the first time frequent blackout complaint early warning appeared, data which is two months before the warning time point is synchronized and previous day's data is synchronized at 0 o'clock everyday, which can meet needs of the data mining.

2.2.2 Establish data mining algorithm of early warning

Due to the fact that 95598 business personnel of State Grid Corporation of China records the failure address of the customer by answering the telephone, the format of the address provided by the fault work order is slightly different from address provided by the power outage information, and the direct use of this address would not meet the demand for statistics on the number of blackouts. Therefore, by referring to the existing Chinese word segmentation matching algorithm based on standard address database, this paper proposes a data mining algorithm that meets the needs of the early warning model to achieve the word segmentation, matching, and collection of the address, and ultimately the statistics of the number of power outages. In the process of data processing, the lexical length is defined by using the length of the field in the data table of the standard address database, the last matched administrative division is adopted to narrow the address range to be matched next. When the address fails to match, the rule tree is considered to achieve the purpose of the next match directly. When all match are completed, the algorithm is terminated and return to match the successful address. Eventually repair address and power outage information within the address format are made uniform and form an early warning data source.

Establish the specific steps by data mining algorithm: ① build a standard database. The comparison table of the administrative area and business area in the State Grid Corporation of 95,598 knowledge base are standardized to form standard database and create data table of address structure; ② Address matching First, the addresses contained in the data of the trouble ticket and the blackout information are sorted out to obtain the writing rules of the address of the daily business work and matching rule; Second, ambiguity table is set to to

solve the problem of ambiguous address at the database level. The ambiguity table is adopted to establish the relation between ambiguous address and standard address. Third, sort out the algorithm steps, the matching rules, ambiguous address processing methods and artificial processing scene are used to determine the algorithm matching process. ③Transit storage Data of automatically match and manual processing will be stored in the early warning database in accordance with the standard format. At the same time, coordinates information, which is obtained in the conversion process through the map interface, will be stored together in the database to be ready for future map presentation. Data mining algorithm flow is shown in Figure 2.

Fig. 2. data mining algorithm flow chart

2.2.3 Set data mining algorithm parameters

The core of this data mining algorithm is Chinese word segmentation, word segmentation algorithm needs to set specific parameters to implement word segmentation in accordance with the specific needs. This feature allows business people to set early warning parameters on their own. Thereinto, configuration parameters include: ①matching rules: Define the matching order of the failure worksheet data and blackout information data; Early warning starting and ending time: Early warning starting and ending time determines the early warning data synchronization and early warning map display. ③Ambiguous table configuration; The ambiguous address of the standard address is configured to improve the matching degree of the address. ④Implementation plan For the specific time when the algorithm is executed each day, the data mining is started at the specified time point, and can also be triggered by the plan to execute directly.

2.3 Principle Of Early Warning Model For Frequent Outages

Early warning maps can show the results of early warning directly and make definition depending on the specific rules provided, the specific contents include:

2.3.1 Warning level definition

Warning level is provided to the business staff to set function. Different warning levels are set, according to the different times of blackouts and shown different levels by different icons. At the same time, generated frequent power outages can also be displayed and set.

2.3.2 The definition of the display way including

Specific presentation information can be defined, including: address, number of blackouts, number of complaints, statistical time. Click Details to view the power failure worksheet information and outage information associated with the site, as well as the frequent outage complaints that have occurred. The specific display style is shown in Figure 3.

Fig.3.map display of Shuang qiao District Chengde City

3 Innovation and application of results

3.1 Early warning model frequent blackouts complaints based on data mining has the following innovations

3.1.1 The innovation of early warning mode of frequent power outage complaints

In the field of early warning of frequent power outages, there is no map to show the form of early warning methods, and no systematic prevention and control measures have been taken. Only after a customer complaint arises, it will be handled passively. The early warning model constructed in this paper starts from the business scenario determining whether frequent blackouts are set up, and normalizes the address of the failure worksheet data and blackout information data by using the Chinese word segmentation algorithm. so that the number of power outages caused by the power supply enterprises can be accurately and efficiently , automatic recorded, and then complete the visual display of early warning information. This result sets a solid foundation for the formation of complaint management and control system of prevention in advance, incident control and post-incident analysis.

3.1.2 The early warning mode combining address normalization with network map has high applicability and a wide range of reference

In this paper, technique solution expressed by work order address standardization method and the combination with network map technology, which is based on standard address library, has broad referential significance. Due to the complexity of Chinese semantics, it is difficult to cut words in the word segmentation process, and it needs a relatively complete vocabulary. In this paper, algorithm, which dynamically defines the maximum length of the maximum matching,

and the matching of the address database, are adopted to optimize and improve the positive maximum matching algorithm. It realizes the normalization and exact matching of work order addresses, and builds a method the standard address library that can be used to the matching of 95588 work order addresses, can be used in word segmentation analysis and application development 95598 and other types of work orders in the future.

3.2 Application results of early warning model for frequent power outage complaints based on data mining

The standardized model of work order address based on the standard address database is adopted to process the failure repair work order and blackout information data at Chengde District in December 2016. It has realized the standardized output of the address, which is accurate to the village, district, company and other common power supply unit, and provides accurate data support for the number of blackouts statistics and frequent power outage complaints analysis. During the period of application, the complaint volume in Chengde District decreases by 30% and the customer satisfaction rate increases by 2.1%, which actually improves customer satisfaction and excellent service level in northern Hebei.

4 Conclusion

This article aims to put forward the service gate forward and reduce the number of complaints effectively. It reflects the continuous self-optimization and self-improvement of power supply enterprises, and is the best manifestation of the service concept of "you are using electricity and I'm working attentively". Early warning for frequent power outage complaint has been very good results in the practical application of the process. It not only reduces the cost of manpower and material

resources in the prevention and control of frequent power outages, and also provides a reference in the complaint verification, rational planning and maintenance of equipment maintenance, and construction of power construction and so on. In the future, we will continue to improve the construction of the standard address database, constantly amend the early warning parameters and improve the accuracy of early warning. We will further improve the quality of service and gradually improve the early warning system to safeguard the sound brand image of SGCC.

References

1. Tao Hongfei, Wang Zheng, Xie Ying Research on frequent power outage control strategy in distribution network [J]. Rural Electrotechnical, 2015 (06) 37-38.
2. Mao Zhongmin, Zhou Xueli, Zhao Huifen, Ma Lei. Construction of Map Public Service Platform Based on WebMap Engine[J]. Computer Technology and Development, 2012(03) 183-186+191.
3. Cheng Changxiu, Yu Bin. A Matching Method for Fuzzy Chinese Address Segmentation Based on Rules[J]. Geography and Geoinformatics, 2011(03) 26-29.
4. Lu Ning. Design and Implementation of Web Map Service Based on WebGIS[D]. China University of Geosciences, 2006.
5. Xu Mingjie, Liang like. Design and Implementation of Spatial Multidimensional Data Model and OLAP[N], Journal of Surveying and Mapping Institute, 2002, 19(2): 124-127.
6. Jiang liangxiao, Cai zhihua, Review and Prospect of Spatial Data Mining[N], computer engineering, 2003, 26(6): 9-11.