Knowledge Discovery and Analysis of ICT Customer Service Data Relying on Multi-frequency Mode Data Mining

Zhongping Xu¹, Tao Zhang¹*, Rongrong Zheng², Xiangna Li³, Dong Li⁴

¹Data consulting business center, Beijing SGITG-Accenture Information Technology Co., Ltd., Beijing, China
²Customer service center, State Grid Information & Telecommunication Co., Ltd., Beijing, China
³Beijing FibrLink Communications Co., Ltd. State Grid Information & Telecommunication Co., Ltd., Beijing, China
⁴Information & Telecommunications Company. State Grid Shandong electric Power Company, Shandong, China

*Corresponding author e-mail: zhangtaofrank@163.com

Abstract. With the rapid development of social economy and technology, ICT often accumulates a large amount of customer service data. How to quickly extract information from it and discover some deep-seated internal laws hidden in business data. This article takes telecom operators as an example, relying on Multi-frequency mode data mining methods, a series of data mining technology application solutions are proposed, and data mining methods are discussed at the same time. Finally, a case is used to prove that data mining is an effective tool for improving the core competitiveness of the business.

Keywords: Multi-frequency Mode Data Mining, ICT Customer Service Data, Knowledge Discovery and Analysis

1. Introduction

With the rapid development of social economy and technology, ICT often accumulates a large amount of customer service data, how to quickly extract information from it, and discover some deep-seated internal laws hidden in business data [1-2]. Telecom operators are representatives of the accumulation of ICT data. With China Telecom's reforms and regulations, the market environment of China's telecommunications industry has undergone fundamental changes. After China's entry into the WTO, telecommunications companies are facing huge challenges and opportunities [3-4]. Some world-class telecommunications giants will gradually penetrate the Chinese telecommunications market through a variety of means, and domestic telecommunications operators are facing a new and more intense market competition environment. Outdated management methods and operating methods have been difficult to adapt to the trend of information society. Market competition is becoming increasingly fierce, customer life cycles are constantly shortening, and the use of advanced management methods and technologies has become the only way for telecommunication companies to solve problems. The decline in telecommunications profits due to increased competition has forced domestic operators to...
also consider seeking new profit models and how to improve their core competitiveness [5-6]. Data mining is an effective tool to achieve this goal.

This paper takes telecom operators as an example, relying on multi-frequency mode data mining methods, and proposes a series of data mining technology application solutions. At the same time, data mining methods are discussed. Finally, a case proves that data mining is the current core competitiveness of the business. Effective tools.

2. Multi-frequency mode data mining

2.1. Modeling rules

We use the multi-frequency pattern technology in data mining to model ICT customer service data. First, we give several definitions of association rules: T has a unique identification TID.

\[
\text{support}(X \Rightarrow Y) = \frac{|T : X \cup YT, T \in D|}{|D|}
\]

\[
\text{confidence}(X \Rightarrow Y) = \frac{|T : X \cup YT, T \in D|}{|T : XT, T \in D|}
\]

2.2. The basic idea of the association algorithm

The specific steps of finding \( L_k \) from \( L_{k-1} \) in the algorithm are as follows:

Connection: Use \( L_{k-1} \) to self-connect to get the candidate \( K \)-item set \( C_k \), \( k \geq 2 \). Let \( I_j \) be two item sets with \( k-1 \) order items, and \( I_{j}[i] \) represents the \( i \)-th item of \( k-1 \) items (\( C_j = 1,2; i = 1,2,\ldots,k-1 \)).

3. Case analysis

After the model is established and verified, there is one main ways to use it. It is to provide market decision-makers as a reference, and he will make suggestions for action plans after viewing and analyzing the model. For example, you can show the clusters detected by the model, the rules contained in the model, or the graph showing the effect of the model to market decision-makers. The model can be used to identify the category of a case, to score an application, etc.

3.1. Customer service sales data set

In view of the above data, multi-frequency mode algorithms can be used to analyze telecommunication bundling sales. It is known that a telecommunications operating company has launched several new services. The company's CRM system sales database has the data set shown in Table 1 below. The transaction mark is represented by the transaction record number, and the item set is represented by the new business and product sold Represented by code.

| Transaction Identifier (TIO) | Sales business items (Items) |
|-----------------------------|-------------------------------|
| 1001                        | NW001,NW003,PH001             |
| 1002                        | MB002,NW003,PH002             |
| 1003                        | NW001,MB002,NW003,PH001       |
| 1004                        | MB002,PH002                  |
| 1005                        | NW002,PH001,MB002             |
| 1006                        | PH001,NW003                  |

3.2. Algorithm to find frequent itemsets

Figure 1 is the process of finding frequent itemsets in D according to the algorithm in the previous section, where the minimum support number (minsupp) is set to 2.
3.3. Association rules generated from frequent items

According to the process of finding frequent items in the previous section, what we get is a frequent set \{NW001, NW003, PH001\}, and its non-empty subsets are: \{NW001\}, \{NW003\}, \{PH001\}, \{NW001, NW003\}, \{NW003, PH001\}, \{NW001, PH001\}. According to formulas (1) and (2), the strong association rules can be output (Table 2):

Table 2. Strong association rule table generated by frequent items

| Association rules             | Credibility | Support |
|------------------------------|-------------|---------|
| NW001 \(\Rightarrow\) NW003 \(\cap\) PH001 | 2/3=67%     | 2/6=33.3% |
| PH001 \(\Rightarrow\) NW001 \(\cap\) PH001 | 2/3=67%     | 2/6=33.3% |
| NW001 \(\cap\) NW003 \(\cap\) PH001 | 2/2=100%    | 2/6=33.3% |
| NW001 \(\cap\) PH001 \(\Rightarrow\) NW003 | 2/2=100%    | 2/6=33.3% |
| NW003 \(\cap\) PH003 \(\Rightarrow\) NW001 | 2/2=100%    | 2/6=33.3% |

With these rules in Table 2, telecommunications companies can formulate reasonable marketing strategies and product cross-bundling plans in future marketing activities, so as to achieve reasonable arrangements for product and service sales. For example, from the rule "NW001 \(\Rightarrow\) NW003\(\cap\)PH001", it can be known that when a telecommunications company provides customers with a new service such as NW001, there is a 67% chance that it can provide customers with another new service of NW003 at the same time and sell PH001 mobile phone services.

Therefore, the telecommunications company can carry out bundling sales activities based on the mined association rules, improve marketing efficiency, increase the output value and added value of a single customer, and strengthen customer stickiness, thereby increasing the overall profit of the

Figure 1. Schematic diagram of the process of finding frequent itemsets
telecommunications company.

4. Conclusion
It is very important to obtain valuable information that is conducive to the operation and decision-making of telecom operators from a large amount of ICT customer service data analysis, so as to selectively bundle sales to customers. Telecommunications companies have collected and stored more and more valuable information about customers, suppliers and business partners. There are many important information hidden behind the surge of data. Data mining technology can discover potentially unknown information in massive data, and can solve the phenomenon of "data explosion but poor knowledge" in telecom companies. As a new business information processing technology, data mining can help telecom operators manage all stages of the customer life cycle. However, the diversity of data mining technology requires telecom operators to carefully select the corresponding model. At the same time, when using data mining technology to analyze customer data, the quality of the data is critical to the effect of the analysis. With the continuous maturity and development of data mining technology, we believe that its role in telecom customer relationship management will be more powerful, which will bring unlimited business opportunities to the telecommunications industry. And for telecom operators to make scientific decisions to provide a strong guarantee.

Acknowledgments
This work was supported by the science and technology project of State Grid headquarters (No.5700-201952259A-0-0-00). It was entitled with ‘Research on key technologies of ICT system fault analysis and auxiliary research based on knowledge discovery’.

References
[1] Ernesto, Aparicio-Puerta, David. liqDB: a small-RNAseq knowledge discovery database for liquid biopsy studies.[J]. Nucleic Acids Research, 2018,4(3):1-8.
[2] Pourghhebleh B, Hayyolalam V, Anvigh A A. Service discovery in the Internet of Things: review of current trends and research challenges[J]. Wireless Networks, 2020,3(1):101-110.
[3] Neely, Benjamin, A. Proteomic Analysis of Urine from California Sea Lions (Zalophus californianus): A Resource for Urinary Biomarker Discovery[J]. Journal of proteome research, 2018,2(1):43-49.
[4] Manol J, Joshi A D, Larson M G, et al. A Discovery-Based Proteomic Analysis Identifies Circulating Growth Differentiation Factor-15/Macrophage Inhibitory Cytokine-1 is Associated with Gastrointestinal and Colorectal Cancer Incidence and Mortality[J]. Gastroenterology, 2018, 154(6):88-96.
[5] Ozuna A, Liberto D, Joyce R M, et al. baerhunter: An R package for the discovery and analysis of expressed non-coding regions in bacterial RNA-seq data[J]. Bioinformatics, 2019, 36(3):11-20.
[6] Namanja A T, Xu J, Wu H, et al. NMR-based fragment screening and lead discovery accelerated by principal component analysis[J]. Journal of Biomolecular NMR, 2019, 73(12):43-50.