

Research on Applications of Big Data in Mobile Communication

Wu Shaohong^{1,*}, Bi Ran¹

¹Department of Important Communications Research, Security Research Institute, China Academy of Information and Communications Technology, Beijing, China, 100191
346591653@163.com

Keywords: Big data, Mobile communication, Hemingway's novels

Abstract. With the rapid development of the mobile internet and the computer ability, the era of big data has come. This paper firstly describes the concept of big data, and then introduces the key technologies of big data in mobile communication network. Then, this paper gives the specific applications of big data in mobile communication, and finally proposes some suggestions on how to make use of the big data technology better to optimize the mobile communication system.

I. Concept of Big Data

Big data is a relatively abstract concept. Currently, the definition of the concept of big data is not uniform. Schonberg known as the "the first man who apply big data in business" points out that the "big" in big data does not refer to the absolute number but the processing mode. We will try to collect comprehensive data, complete data and synthetic data, at the same time, the use of mathematical methods to analyze and its modeling, mining behind the relationship, thus predicting the probability of event happening behind. Tu Zipei, a famous information management expert of our country, believes that the "big" in big data an imaginary reference in his book big data. It refers to the human ability to find large value data contained in the big data. Big data is often equated with the concept of "massive data" and "ultra large scale data" terms of formulation. However, its connotation greatly exceeded the traditional data form, also beyond the processing capacity of the existing techniques. The most representative definition of many big data definitions is the 3V definition. The first V means volume, namely the data volume is huge; the second V means variety, namely there are a lot of data types in big data, such as text, images, audio, video and so on; The third V means volume velocity, that is the fast processing speed of big data. In addition, International Data Corporation (IDC) believes that big data should also have the high commercial value. IBM believes big data should be veracity. In short, the definition of big data concept can be based on 3V or 4V or 5V.

II. Key Technologies of Big Data in Mobile Communication Network

Clustering is a process of grouping physical or abstract data objects. Clustering analysis is widely used in data mining, statistics, machine learning, pattern recognition, biology, and spatial database technology and e-commerce applications. Data mining clustering methods are mainly divided into the hierarchical method, density-based method, and grid-based method and model-based method.

Decision tree is a flow chart like tree structure. The each internal node of the tree represents a test on an attribute, its branches on behalf of test results for each, and each leaf node of the tree represents a category, the tree top node is the root node is the beginning of the decision tree. An

example consists of a set of values and a predicate goal value of the target. Decision tree algorithm is a method of approximating discrete value function, and it is also an induction algorithm. It uses "top-down, divide and conquer method of the search space is divided into several disjoint subsets. It is usually used to form classification and predictive model to preprocess the unknown data.

The CRISP-DM is the standard of the current authority in data mining models. CRISP-DM data mining process is divided into six stages: business understanding, data understanding, data preparation, model building, model evaluation and results publishing. Business understanding is an understanding of business operations, business processes and industry background. Data understanding is an understanding of the existing enterprise application systems. Data preparation is to remove a large number of data from the enterprise to explore issues related to the model data subset. Modeling is based on the understanding of business issues, on the basis of data preparation. It chooses a more practical mining model, the conclusion of the formation of mining. Evaluation is in practice, the results of the excavation test, if you achieve the desired results, you can publish the results. But in the actual project, the data understanding, data preparation, modeling, and evaluation in the CRISP-DM model are not a one-way operation, but a process full of repeated adjustments, and constantly revise.

III. Applications of Big Data in Mobile Communication

A: Application in Network Optimization

Network optimization is the key to network analysis and problem location. The network needs to solve the problem is mainly interference, loss of words, traffic imbalance and switch fault four aspects. The purpose of optimization is to make the wireless signal continue to be well covered, while reducing the area of the non-dominant area, reducing interference, loss of speech, traffic imbalance and switching failure four aspects of the problem. Network optimization includes the optimization of preparation, data collection, problem analysis, optimization and adjustment of the implementation, the extraction of data confirms the five parts. The data acquisition and analysis, optimization and adjustment need and network optimization after repeated according to the wireless network requirement until the wireless network to meet the optimization goal to meet the requirements of the key network metrics. We adjust the background parameters according to the adjustment plan so as to ensure that the next step in the business optimization of wireless signal distribution is normal, to optimize the work to lay a good foundation. Team network optimization personnel adjust the unreasonable region of the antenna radio frequency and modify the background parameters according to the adjustment plan to optimize the communication network.

B: Application in Marketing Management

Big data analysis cannot be separated from the data quality and data management. High quality data and effective data management is a prerequisite for the effectiveness of large data. Therefore, we need to clean and preprocess data firstly to optimize the marketing management. Data cleaning and pretreatment are the data preparation work before modeling. Data cleaning and pretreatment of assurance data modeling is correct and effective, on the other hand, through the adjustment of the format and content of the data, the model was more accurate and efficient, to ensure data quality and availability. Before the model is set up, the sample data is divided into two parts: one part is used for modeling, and the other part is used to modify and test the model. After the model is set up, it needs to use a large amount of data to test and verify the model repeatedly. The model is validated by the actual data, and the model can be trusted. After cleaning and pretreatment of the data of marketing management, the data quality is guaranteed, so the customer can be subdivided, and the customer group analysis model is established. As mentioned in the CRISP-DM model, there is no

strict order in the six stages of data mining, and it is possible to modify the data alternately in several stages.

C: Application in Fee Computing

With the rapid growth of mobile users and the volume of business, the wandering fee of mobile phone needing the fee computing system must deal with is increasing fast. The voice roaming words alone has reached the number of daily average of nearly 200 million in China. During the holidays, especially during the Spring Festival, the volume will have a sharp increase, or even doubled. The volume in next few years is expected at an average annual rate of about three percent. The existing software and hardware architecture has become more and more difficult to support, and many new business systems also have a high demand for the massive data processing capacity of the system. Through the analysis, it is found that the performance of the system is mainly in the efficiency of database reading and writing. The main reason is that the data is processed and the processing time is large. After several projects and found that the above mentioned two strategies, whether it is optimizing software or hardware upgrade just to satisfy short-term needs and under the existing structure, upgrade costs high, only hardware upgrade cost often in the millions. Integrated technical and economic factors can improve the processing performance.

IV. Suggestions on Applications of Big Data in Mobile Communication

A: Develop Storage Technology

Along with the advent of the era of big data, the mobile communication enterprises are facing massive information storage requirements. The urgent problem is to enhance the data storage capacity, because only when the data is properly stored, it has the possibility to further tap the potential value. Enhancing the data storage capacity for operators provides new business opportunities. Major telecom operators to use their own advantages have launched the construction of data centers and cloud computing center of the grand plan. Aiming at the demand of large data capacity, storage virtualization is now increase the capacity efficiency is the most important the most effective solutions, it is lack of the ability of these existing storage system to expand the storage efficiency of the automatic hierarchical and streamlined configuration tool. Virtualization supports hiding details and complexity, and ultimately makes storage services flexible, scalable and manageable. Have the virtual storage, can be derived from the internal and external and the multivendor storage of structured and unstructured data file and block content and storage of all data types, and integrated into a single storage platform. The data virtualization can change its dynamic capacity and reduce the consumption of storage resources.

B: Clear Applications Target

Dramatic changes have taken place in the current telecom operators. They are facing the survival and development environment changing. The market saturation ascension leads to more intense competition between each communication system. The homogeneous competition trend has become increasingly prominent in micro channel for business began to erode the traditional telecom operators in the business field. Telecom operators feel the pressure of unprecedented theoretical research, the transformation of the development of telecommunications operators to consider the main problem. The data flow through the mobile phone terminals, operators and application providers and eventually enter the people's attention in the big data era. The data as a core asset is a profound impact on the development of all enterprises. The boundary of the telecom operators and service providers to become blurred, through cross-border multi technology integration to build the integration of multi-service platform, which is full of advantages of large data reorganization, conform to the change management. It grasps the lifeline of enterprise development. Telecom

operators provide the Internet all the data channel, the natural advantages and has a huge user base, if we can make full use of these advantages, telecom operators will become the most powerful competitors in the era of big data. We need to actively deploy data mining analysis business, extract high value business data in large data traffic to increase the business revenue.

C: Protect User's Privacy

With the generation, storage, analysis of the amount of data increasing, how to protect sensitive data or protect the privacy of the data is becoming increasingly important. Data theft is not only exposed the personal information of consumers and enterprises confidential information, but also is exposed the national security secrets. In view of the serious data stolen events unabated, the technical and policy tools solve data security problems will become reasonable. The release of data generated from the data and information collected data, analysis and application, data copyright involves every aspect, our country in the laws and regulations still remain blank to some extent. The safety of the data management needs the government and laws to regulate. Enterprises need to pay attention to the safety of data of user's behavior in the process of the data of analysis to prevent the leakage of data. Any data leakage will do harm to the good reputation of the telecom operators.

V. Conclusion

As the core status of the communications industry chain, telecom operators should realize that the data is the core strengths to carry out mobile Internet business assets. The mobile communication system can be more popular by using the big data technology. Telecom operators have an unparalleled advantage in the era of big data. The value of large data mining will bring huge returns.

References

- [1] Zhang Jun, Research on Key Technologies of Big Data in Mobile Communication Network [J] Telecommunications Network Technology, 2014 (4):10-12
- [2] Yu Haibo, Application of Big Data in Telecommunication Network Optimization of Mobile Communication [J] Guangxi Communication Technology, 2014 (4):8-11
- [3] Zheng Shuang, Luo Chunling, Huang Yang, Research of Mobile Communication Industry Micro Marketing in the Age of Big Data [J] Guangxi Communication Technology, 2015(1):6-9
- [4] Liu Zhihui, Zhang Quanling, Research Overview of Big Data Technology [J] Journal of Zhejiang University (Engineering Science), 2014, 48 (6): 957-972