

Research on Internet of Things Data Mining Based on Cloud Computing Platform

Enhong Liao

Guangdong Engineering Polytechnic College 510520

Abstract: With the rapid development of the current society, the value of the information industry in the development of the national economy is becoming more important. The Internet of things is one of the most important technical means. As an important component in the Internet of Things, data mining technology plays an important complementary role. This paper briefly analyzes the basic connotation of cloud computing and Internet of Things (IoT), expounds the typical features of data mining technology, and focuses on the research of IoT data mining based on cloud computing platform. It has high practical value through experimental testing.

1. Introduction

Internet of things technology has made tremendous progress, and has been extremely widely used in people's daily life. A large number of IoT applications are beginning to take people's perspective, but at the same time, the current application of IoT technology also raise more stringent target requirements. The face of the current challenges to the Internet of Things data mining challenges, through the introduction of cloud computing technology can promote the development of the Internet of Things which has a more powerful IT infrastructure support capabilities, and the ability to analyze data mining and platform to open up the ability to expand a great help. Therefore, cloud computing platform for Internet of Things data mining related research has great practical significance.

2. Cloud Computing and Internet of Things Analysis

2.1 Cloud Computing

Cloud computing is an add-on application and delivery model based on Internet-related services that typically involve the use of the Internet to provide dynamic, easy-to-extend, often virtualized resources. At present, the mainstream view of the definition of cloud computing is man-made. In this case, a mode of charging based on the amount of usage can provide a convenient, efficient and network-based access mode according to the demand and enter the pool of configurable computing resources, and resources can be distinguished efficiently, requiring minimal management effort or minimum interaction with service providers.

2.2 Internet of Things

Internet of things is not only the core content of the new generation of information technology, but also a major stage of development in the information age. Understand literally, Internet of Things and the interlinkages between things, its content mainly includes two aspects: (1) Internet of things is still the Internet as the core and foundation, and is based on the Internet extends and extends the network form; (2) Internet of Things client directly extended to any two items to achieve the purpose of



information exchange and communication, that is, the interrelationship of things. The Internet of Things can be applied to the field of network convergence by means of intelligent perception, technical identification and applicable calculation, and thus the Internet of Things is also considered as the third wave of the development of the information industry after taking over the computer and the Internet. As an important extension of Internet applications, the Internet of Things has more than just a purely Internet connection, but also an effective connection between business and applications. Therefore, under the current development of Internet of Things technology, its core is always based on application innovation and user experience as the core.

3. Data Mining Technology

3.1 Connotation Definition

Although the history of data mining is not very long, its value of human society is becoming more important. Considering that data mining itself is an interdisciplinary field, people in all fields must understand that there will be different differences. Therefore, there is not a unified definition standard for this technology at present, and different persons also put forward completely different definition statements based on their own research content and application objects. At present, the more mainstream view is that data mining is based on large-scale data-based exploration and simple correlation model selection of an efficient means of measures, there are researchers of man-made data mining technology is the use of a number of significant differences exist in the technical means. To find a significant new relevances to the massive data and information, new model and the development of new trends [1].

The currently accepted definition of data mining specifically includes the following aspects: (1) The total amount of data sources must be sufficient, and true and reliable, it should be noted that the actual data may not be complete or even including the presence of noise; (2) the information obtained through data mining or the content of knowledge has very special value meanings for the users; (3) the finally found knowledge elements are acceptable and understood that can provide certain reference support for the final decision.

3.2 Challenges

According to IoT data characteristics, the characteristics of IoT data now challenge the data mining technology as follows:

(1) A large-scale IoT data information is stored in different locations respectively. Therefore, it is more difficult to adopt centralized data mining for distributed data mining.

(2) Because IoT data information is often faced with a large number of external factors, such as data security, privacy, legal restrictions and other issues. So it cannot stored all the data and information in the same data warehouse.

(3) The content of the data contained in the Internet of Things is very complicated. In addition, there are many sensor nodes at the same time and must be dealt with in time. Usually, the structure of the central processing unit is applied. Therefore, the hardware requirements of the central node will be raised higher.

(4) Considering the limitation of the limited resources of nodes, storing all the data information about the central node can not realize the optimization and application of expensive resources. In most cases, the central node does not need to use all the data information, but need to expand the parameters in advance prediction and evaluation, which can be based on distributed nodes for each of the underlying data contents early take measures to deal with, and then applied to the information sent to the receiving staff.

Based on the above analysis, it can be found that there are still many shortcomings in the current techniques and forms of data mining for the Internet of Things, and more research is still needed to develop better solutions.

4. Internet of Things Data Mining Based on Cloud Computing

The application of IOT data mining technology in cloud computing platform mainly includes the following four items:

4.1 Perception of Things Layer

The value of this layer is mainly reflected on the deployment of a large number of acquisition nodes based on target area and the use of such nodes to obtain IoT data information about sensors, cameras or other equipment. The corresponding data and information contents can communicate with each other within the Internet of Things perception layer, that is, wireless sensor networks are generated, the networks are aggregated on the nodes, and then the data and information contents are aggregated and stored, send to cloud platform data center.

4.2 Transport Layer

The value of this layer is mainly the sensor networks, wired networks, wireless networks and other network forms converge in the same overall, and in order to achieve efficient and complete data transmission of information can be very convenient to promote the perception of data information is sent to the cloud computing data center, and ultimately to achieve more complete data information intercommunication; to promote a variety of different monitoring equipment can achieve data networking transmission, so that the IoT monitoring equipment networks data transmission and efficient information [2].

4.3 Data Layer

This layer has a crucial role in the entire IoT data mining platform. In view of how to fully understand the heterogeneity and massiveness of IoT data, how to solve the correlation stored in the IoT The data information content will also directly affect the feasibility and application performance of the Internet of Things data information mining platform. Data layer contains two key modules specifically, namely data source conversion module and distributed storage module. The former module is mostly used in heterogeneous data conversion to the Internet of Things, and the latter module are even more fully integrated with the cloud computing platform file HDFS. The application is distributed network data storage to achieve large-scale networking record of information [3].

In the Internet of things, the difference between the object itself will be represented by a variety of data types, and even the same type of data will be used to represent different data sources. Therefore, the value of the data sources onversion device is mainly reflected on the difference between IoT data. The handling of fabric issues not only helps to ensure better integrity of the data store, but it also helps smooth the data mining effort. The function value of the data sources conversion module can be basically compared with the interfaces of the monitoring devices in the data layer and the perceptual layer. At the same time, the content stored in the distributed storage module are caused to be complete by decoding the data packets and corresponding data models information. Data conversion devices can facilitate the conversion to many types of data onto PML format, so the file data distributed among multiple nodes are PML type data.

As a result, the concept of PML is introduced as a data type. PML can describe natural objects of a widely used form and is a new language evolved from the basis of XML. Have similar ideological content [4]. With the emergence of PML can provide accurate information about the relevant items, and enhance the exchange of information on the items. If the information acquired by the IoT nodes is transmitted, the PML expansion model can be applied to the storage when it is stored. The content of the information contained in the IML node mainly includes the attribute information about the object itself, the location information, the environment in which a single object is located of the specific information, or even the history of the object content of the information in the above information content can be combined to make a more accurate description of the article.

4.4 Data Mining Service Layer

In this layer, its focus on the content of the data pre-prepared modules, data mining engine modules and user modules in three parts. The preliminary preparation module mainly aims at the content of data and information such as finishing, transformation and restriction. The data mining engine module focuses on the evaluation and analysis of data mining algorithms and modes. The user module specifically represents the visual representation of data mining knowledge content [5]. According to the different mining ways of knowledge, the functions that will be applied to the data mining engine module mainly include the characteristics differentiation, association, type aggregation, spectacle, trend prediction, deviation analysis, analogy analysis and so on. The most important point to provide the above-mentioned functions is that the algorithm set in the data mining engine module simultaneously provides a plurality of different function calculation methods. On the cloud computing platform, data mining calculation methods need to be improved on the data mining methods used in the past, which is to carry out the algorithm application and expand the optimization process. The user module can be in direct contact with the application personnel of the whole IoT data mining platform, and therefore, the user module needs to have a good friendliness feature. The user can work through the interface to perform card data mining work in order to gain understanding and support [6]. In order to promote the platform port attributes effectively, an open connection channel can be added at the bottom of the user service module to facilitate the retrieval and application of functions of the IoT data mining platform by a third party so as to ensure that the IoT application can achieve more richness degree.

5. Experimental Test

5.1 Work Process

The process of Internet of Things data mining is shown in Figure 1 below. After the user sends out a request for data mining, the main control node will make a judgment after receiving the request and can perform this task. At the same time, the task of feedback on the user can be carried out. If successful, the main control node retrieves and applies the data mining calculation method required by the user in the storage module based on the data mining algorithm, and then the relevant data information contents in the HDFS files storage system will also carry out work such as data protocol [7]. After that, we can carry out distributed data mining based on the results of the algorithm calculation. In carrying out this work, we mainly apply the master / slave structure relationship, and after the master node completes the task of data mining division, it further conducts to the required completion of the actual work node, the corresponding node is mainly related to data mining data processing.

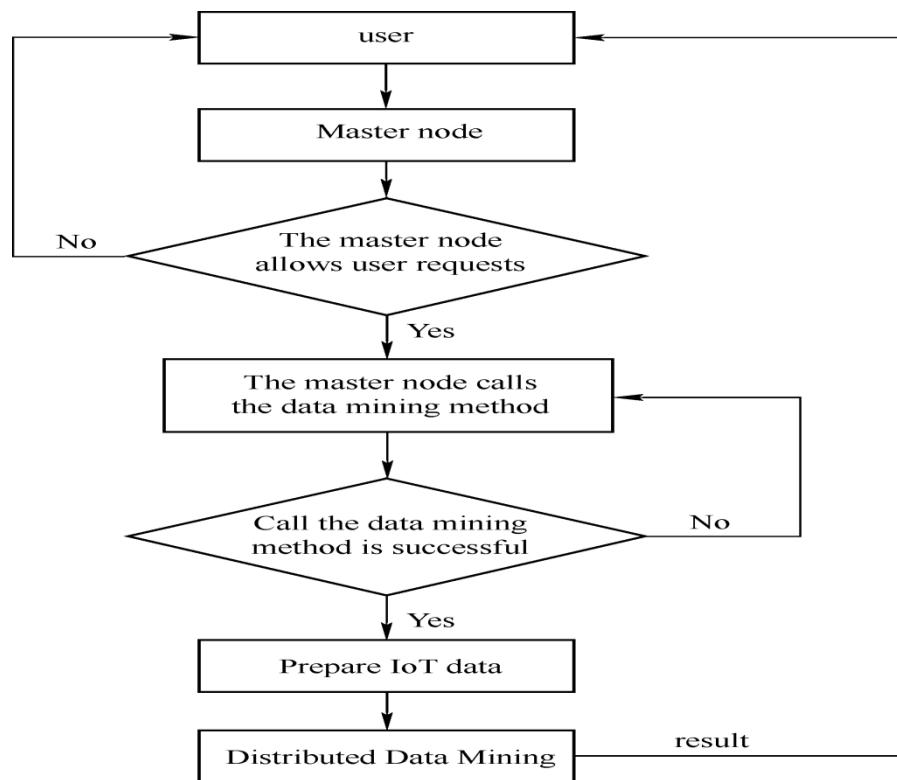


Figure 1 IoT data acquisition flow chart

5.2 Experimental Test

Build a cloud computing platform, and at the same time implement the distributed storage through the data information transformed by the data conversion device into the PML format, and analyze the feasibility of the cloud computing platform through the data mining calculation method. After completing the configuration work on the cloud computing platform, a set of experimental data applied to the association rules algorithm is selected and the program written in C ++ code is converted into the PML format by the keyword search method and put into the cloud of the HDFS instruction computing platform for distributed storage [8]. After running the system, we can get the result and check whether we find all the frequent item sets in the experimental data set.

6. Conclusion

All in all, the research focused on data mining for IoT based on cloud computing platforms. In view of the characteristics of the IoT itself also indicates many difficulties and challenges on the development of IoT data mining work, in order to effectively solve the problems related to the difficulty, this article carries / carried out an in-depth analysis and discussion on the Internet of Things data mining and cloud computing technology, put forward the use of cloud computing platform to carry out distributed data mining research, and through experimental testing confirmed the feasibility of this approach.

Author:

Liao Enhong, Guangdong Engineering Polytechnic, vice president of information engineering, associate professor, a construction division, mainly engaged in computer information-related professional teaching management, cloud computing big data and Internet of Things related research, information systems engineering project management.

References:

- [1] Feng Huining. Hierarchical Modular Modeling and Design for Multiple Streams in Cloud

- Computing [J]. Systems Engineering Theory and Practice, 2013,33 (6).
- [2] Zhang Yongping, Zhang Gongxuan, Zhu Zhaomeng et al. Cloud Acceleration Method for Compressed Sensing in Internet of Things [J]. Journal of University of Electronic Science and Technology of China, 2014, (3).
 - [3] Dou Xin.Design of integrated information service platform for logistics park integrating cloud computing and IOT technology [J]. Modern Electronics Technique, 2017,40 (9).
 - [4] Zhang Wenli, Guo Bing, Shen Yan, et al.Study on computing migration of intelligent mobile terminal [J].Acta Ch Computers, 2016,39 (5).
 - [5] Ma Liangliang.Analysis of data mining patterns of Internet of Things under cloud computing [J].Electronic Testing, 2017, (4).
 - [6] Zhang Jiandong.Study on the Convergence and Development of Cloud Computing and Internet of Things [J].Journal of Information and Computer, 2017, (18).
 - [7] Lv Hai-Dong, GE Ri-Bo.Application Development of Distributed IntelliSense IOT for Cloud Platform [J].Computer Measurement and Control, 2016,24 (7).
 - [8] Zhang Fucui.Application Analysis of Cloud Computing and Internet of Things in Building Smart Cities [J]. Information & Communications, 2017, (5).