Fusion Application of Big Data and Cloud Computing In the Internet of Things

Yongjun Qi\textsuperscript{1} and Haiyan Wu\textsuperscript{2,*}

\textsuperscript{1}Faculty of Megadate and Computing, Guangdong Baiyun University Guangzhou, China
\textsuperscript{2}Management School, South China Business College Guangdong University of Foreign Studies, Guangzhou, China

*Corresponding author e-mail: wuhaiyan@gdufs.edu.cn

Abstract. The Internet of Things can be simply understood as the mode of connecting things in the Internet environment. It is mainly the organic combination of computer, Internet technology and other information technology means. At the same time, the organic combination of the Internet of Things, big data, cloud computing and other advanced information thinking and computing methods is the direction of the future smart city construction. In the experiment, due to the characteristics of real-time and massive data of the Internet of Things, the fusion processing of big data cannot be carried out well. Based on the Internet of Things technology, cloud computing technology and fusion algorithm are selected for research. Experimental data show that the number of four types of nodes is set as 30, and the system has 100 nodes in total, and the length of one rotation of the system is 50s. The ordinary Leach algorithm is used first, and then the improved Leach algorithm is used in the same environment. The experimental results show that there are four different types of nodes in the monitoring environment obtained by the test, and the initial data transmission rates of each node are 60, 90, 110 and 80 respectively. Therefore, the adjusted and improved LEACH algorithm can allocate the required data transmission time according to the different failure levels of the data. If the aging level is higher, the data transmission time of node data will be relatively longer. The improved strategy is in line with the characteristics of the Internet of Things and has the feasibility to complete the experiment.

Keywords: Internet of Things Technology, Big Data Technology, Cloud Computing Technology

1. Introduction

Big data, cloud computing and the Internet of Things are inextricably linked. Moreover, the physical world and the information world have been effectively integrated, and people cannot live without information networks. Therefore, it is imperative to build a smart city in the future [1, 2]. In the construction of smart cities, the Internet of Things is a very important part. To realize the information
interaction between things and people, big data and cloud computing are needed as the support. Therefore, it can be said that big data and cloud computing are the important foundation supporting the development of the Internet of Things [3, 4].

Konikov et al. believe that the development of cloud computing is the driving force for the rapid development of data centers. People store more and more data, such as pictures and videos, in data centers. Therefore, the number of storage devices needed in data centers keeps increasing, which requires the expansion of the scale of data centers [5]. Chen et al. believed that the upper business carried by the data center was also changing, such as Map-Reduce, distributed storage, virtualization, etc., and these upper businesses promoted a large amount of communication between servers within the data center, and the traditional structure of the data center could no longer meet the requirements in this respect [6].

The distribution of the cloud computing will be distributed computing jobs, will be assigned to a task more child to operate on a computer, thus avoids the problem of "overheating" part of the node, distributed processing tasks at the same time reduce the workload of each node, reduces the requirements on the properties of the node, you can use the mass of cheap computing equipment complex computing task [7, 8]. With the progress of Internet technology and technology, as well as the rise of a series of new applications such as intelligent transportation, smart home and intelligent logistics, the Internet of Things is now the pillar of the national economy in the future [9]. If both big data and cloud computing are combined, there are both opportunities and challenges for the Internet of Things. Therefore, this paper introduces the application of big data and cloud computing in the Internet of Things, and explains the relationship and characteristics of the three, hoping to provide a better foundation for the future development of the Internet of Things [10].

2. Research on the Convergence and Application of Big Data and Cloud Computing in the Internet of Things

2.1. Application of Big Data and Cloud Computing in the Internet of Things

There is an interaction between big data and the Internet of Things. First of all, the Internet of Things and the mobile Internet are the sources of big data, and big data provides effective information and useful analysis results for the Internet of Things by collecting and analyzing user behaviors, thereby obtaining value. Cloud computing is the foundation of big data applications. Without cloud computing facilities, big data computing cannot be realized. If the Internet is compared to people, then the Internet of Things is its perception and motor nervous system. The key role is to analyze and count the data received by the big data layer of the Internet. The Internet of Things can realize the communication between things and things and between people and things. Cloud computing is the central nervous function of the Internet of Things. Under normal conditions, Internet users communicate with cloud computing based on the sensors and terminal systems of the Internet of Things, and aggregate data and receive services for them.

2.2. Application of Cloud Computing and Internet of Things Integration

We describe a range of applications that could be widely used or significantly improved in the future due to cloud computing and the Internet of Things.

(1) Medical care

The Internet of Things and Internet technologies will be able to contribute to healthcare based on environmental assistance and telemedicine. Nowadays, smart devices and Internet devices, for example, play a prominent role in the health care system due to their powerful characteristics, making efficient, high-quality, reasonably priced services widely applied in the medical field. At present, the mass data formed in the common medical field must need other sensor data analysis and processing, and give proper management. Give cloud solutions, can make more abstract specific technical details, can solve the professional knowledge required for the application of the technology infrastructure, can form an effective medical sensor for the solution of the huge amounts of data in a reasonable manner
to use cloud technology to medical sensor data to provide efficient and reasonable solution, can let he further mobile devices for the exchange of information, makes the safety of medical data accuracy.

(2) Smart city
The Internet of Things provides general-purpose middleware for the smart city services of the future, retrieving information from a variety of heterogeneous sensing infrastructures as well as from all types of geo-location and cloud technologies, such as 3D representations with RFID sensors and geo-tags. You can access it. And it presents information in a uniform way. Many recent solutions have proposed ways to use cloud architecture to detect, connect, and integrate sensors and actuators, creating a platform that provides and supports ubiquitous smart city real-time applications. And the framework can include a sensor platform (for detection and driving) and a cloud platform that can automatically manage, analyze and control large amounts of data from real-world devices. This high-level service model hides the complexity of the underlying cloud computing infrastructure, but provides security, heterogeneity, interoperability, scalability, and responsiveness to the public cloud. Make it complicated. In addition, cloud-based platforms facilitate the development and delivery of cloud plug-ins for third parties to connect the Internet of Things to any device.

(3) Smart energy and smart grid
Both the Internet of Things and cloud computing can achieve effective coordination under heterogeneous local conditions and a wide range of environments to complete intelligent management tasks, such as energy distribution and energy consumption. And under some conditions, different types of nodes, the information collected can provide lighting for opportunities. Such nodes have discovery, processing and monitoring functions, but their scope is very limited. So computing tasks need to be appropriately distributed between nodes or within the required cloud to allow for more complex and effective planning. Another situation is to integrate cloud data into the system to achieve energy substitution and compatibility, and provide self-regulation, interoperability, user participation, power quality optimization, distributed generation and demand response, which can solve some usage problems. The integration of cloud computing platforms into the Internet of Things scene has caused people to worry about the security and privacy issues of smart grid software. In order to fully exploit the potential of these applications, these issues need to be completely resolved. Customers need to build trust and share data so that they can improve and enhance the services provided.

2.3. Fuzzy Data Fusion Algorithm
This paper introduces the algorithm in detail through several main functions. The main functions of the fuzzy algorithm include fusion function, admissibility function and fuzzy measure.

The first is the notion of fusion functions. The fusion function mainly solves the problem of how and what algorithm the data will be fused. The simplest methods include the average method, the maximum (small) value method, the middle value method, etc. Of course, there are also more complex algorithms, which are determined according to the specific requirements of fusion. Suppose there are N sensor nodes in the multi-sensor data fusion system, and their output data are X1, X2, and..., Xn, then the fusion function of the system can be expressed as:

\[ F(x_1, x_2, ..., x_n) = y \]  

(1)

In the above formula, F represents the fusion function, and y represents the result of data fusion of n nodes. The fusion function should have the properties of commutative and idempotent functions.

The next thing I want to introduce you to be the concept of an admissibility function. As the name implies, the tolerance function describes the degree to which data collected by two or more sensor nodes can be fused, which means the similarity degree of node data. The higher the similarity degree of data and the closer the data are, the higher the values of the tolerance function will be.

The value of the permissible function is specified on the interval \([0, 1]\).

When multiple data are waiting for fusion, the tolerance function is defined as follows:
In the above formula, $R$ represents the tolerance function, $R(X_i, x_j)$ represents the tolerance result of two sensor nodes, $R(x_1, x_2, ..., x_n)$ represents the overall tolerance result of the data of $N$ sensor nodes, thus it can be seen that the value of the tolerance of multiple sensor nodes is obtained by taking the minimum value after comparing the tolerance of data of two nodes.

3. Experiments on the Fusion of Big Data and Cloud Computing in the Internet of Things

3.1. Experimental Setup

In the face of real-time and massive data characteristics of the Internet of Things, big data fusion processing cannot be performed well. Based on the Internet of Things technology, cloud computing technology and fusion algorithm are selected for research, and solutions are proposed, so that the improved data fusion can be well applied in the Internet of Things environment. The experimental software environment was configured as Windows XP operating system, and NS2 network simulation software was run. The properties of the improved LEACH algorithm and the improved fuzzy algorithm were simulated to test the practicability of the algorithm.

3.2. Experimental Purpose

Different types of data that need to be integrated should be distinguished in IoT technology, that is, the importance of all kinds of data should be different, and more important data should reach the fusion node faster for data fusion, so as to conform to the characteristics of the Internet of Things. However, the LEACH algorithm's average allocation of data transmission time obviously fails to meet the requirements of the Internet of Things, the experiment conducted further research on LEACH algorithm and proposed an improved scheme based on the original algorithm, so that the improved algorithm can be more adapted to the environment of the Internet of Things.

4. Discussion on the Integration and Application of Big Data and Cloud Computing in the Internet of Things

(1) Set the number of four types of nodes to be 30, the system has a total of 100 nodes, and the time length of one cycle of the system is 50s. First use the ordinary LEACH algorithm, and then use the improved LEACH algorithm under the same environment. The experiment first performs statistics on the time between when the data starts to be transmitted by the node and then is completely accepted by the cluster head node, and then the optimized LEACH The algorithm is simulated, and the heterogeneity of the IoT environment needs to be highlighted in the simulation process. For better analysis, we assume that there are four different types of nodes in the monitoring environment. The initial attribute values of different types of nodes are shown in Table 1 and Figure 1:

| Node type | Aging grade | Data sending rate (Kb/s) |
|-----------|-------------|-------------------------|
| A         | 4           | 60                      |
| B         | 3           | 90                      |
| C         | 2           | 110                     |
| D         | 2           | 80                      |

Table 1. The initial state of the four types of nodes
According to the above discussion and analysis, the nodes with the highest demand for real-time performance and the highest aging level of the system are type A nodes, the lowest aging level nodes are type D nodes, and the nodes with the same aging level are type B nodes and type C nodes. Compared with class C nodes, the data sending rate of class B nodes is slightly lower. The ultimate goal of the experiment is to match the various data transmission time with the node's aging level. The higher the aging level, the more data transmission time will be. When the same type and aging level are also the same, the data is sent the result of the rate is Table 2 Figure 2:

| Node type | Aging grade | Data sending rate (Kb/s) |
|-----------|-------------|------------------------|
| A         | 1           | 80                     |
| B         | 1           | 100                    |
| C         | 1           | 120                    |
| D         | 1           | 150                    |

5. Conclusion
Nowadays, the information network has spread throughout the daily life of human society, and the construction of smart city and smart life in the future cannot be separated from the Internet of Things, and the development of the Internet of Things is more inseparable from the support of big data and cloud computing. Therefore, the organic combination of the information world and the physical world...
should be achieved. Based on Internet of Things technology, cloud computing technology and fusion algorithm are selected for the experiment. The experimental results show that the initial data transmission rates of four different types of nodes in the monitored environment are 60, 90, 110 and 80 respectively.

Acknowledgments
Supported by (1) the Characteristic Innovation Project in Natural Science of Department of Education of Guangdong Province (No. 2018KTSCX256). (2) The Project in Key Fields of New Generation Technology of Department of Education of Guangdong Province (No. 2020ZDZX3001).

References
[1] Majhi S K. Special Issue on Convergence of Information Systems with Cloud Computing, Big Data, and IoT. 2018, 8(1): vii-viii.
[2] Mishra B S P, Das H, Dehuri S, et al. Fog Assisted Cloud Computing in Era of Big Data and Internet-of-Things: Systems, Architectures, and Applications. 2018, 10.1007/978-3-319-73676-1(Chapter 14):367-394.
[3] Sanagavarapu K C. A Review on Cloud Computing Services and Convergence among Cloud Computing and Big Data. International Journal of Engineering Technologies and Management Research, 2020, 6(9):84-91.
[4] Mavromoustakis C X, Mastorakis G, Dobre C. [Studies in Big Data] Advances in Mobile Cloud Computing and Big Data in the 5G Era Volume 22 || Handling Big Data in the Era of Internet of Things (IoT). 2017, 10.1007/978-3-319-45145-9(Chapter 1):3-22.
[5] Konikov A, Sudakova A. Exploring the possibilities of using machine vision in the construction industry. IOP Conference Series: Materials Science and Engineering, 2020, 1001(1):012046 (6pp).
[6] Chen Y. The Ccms based on internet of things and cloud computing in the age of big data. Revista de la Facultad de Ingenieria, 2017, 32(13):106-111.
[7] Mavromoustakis C X, Mastorakis G, Dobre C. Big Data and Cloud Computing: A Survey of the State-of-the-Art and Research Challenges. 2017, 10.1007/978-3-319-45145-9(Chapter 2):23-41.
[8] Zhang T, Wu X. Research on intelligent logistics development model based on internet of things and cloud computing in big data age. Revista de la Facultad de Ingenieria, 2017, 32(6):341-346.
[9] Crespo-Perez G. Convergence Of Cloud Computing, Internet Of Things, And Machine Learning: The Future Of Decision Support Systems. International Journal of Scientific & Technology Research, 2017, 6(7):131-136.
[10] Yi S. Research on the development mode of intelligent logistics based on internet of things and cloud computing in big data era. Revista de la Facultad de Ingenieria, 2017, 32(16):829-834.