Research on the Application of Big Data Technology in Network Security Analysis

Zhihui Liu

Xi'an International University, Shaanxi, Xi'an, 710077

Keywords: Network Security Analysis; Big Data Technology; Application

Abstract: As the development of Internet, cloud computing technology and big data technology have become more and more popular, network security has become a hot issue of the public at present; it could not only influence the daily production and life of people, but also cause the social crisis of information leakage, even the leakage of national secrets. At present, the network security analysis that combines the big data technology has reached a new stage of development, so how to analyze the specific role of big data technology in network security analysis, and how to truly integrate the big data technology into the work of network security analysis have become a breakthrough in the current network security management work. Based on above, this paper researches the application of big data technology in network security analysis.

As the constant development of the network technology, the Internet has had a great impact on people's life and the way of production, and has become an indispensable medium in the development of society. However, in the process of enjoying the convenience brought by the Internet, the problem of network security has also caused people's deep thinking. From the perspective of its influence, network security could not only influence the security of people's property and confidential information, but also threaten the information security of the enterprise, even threaten the security of a nation; therefore, it has become one of the hot issues of all walks of life. Big data refers to the acquirement, management and collection of data in certain time and scope, so as to obtain comprehensive and diversified data resource. In the process of using the network, a large amount of confidential data is stored or transferred to the huge database that named cloud storage, so the analysis of network security can be done from the perspective of big data; this is the theoretical basis of this paper.

1. Advantages of big data technology in the application of network security analysis

Big data technology plays a very important role in network security analysis; the most obvious role is the function of integrating the data for network security analysis better. Specifically, large data can achieve large capacity information analysis function high efficiently and economically, and it can meet the processing and security requirements of network security analysis. In the network security analysis, big data technology can integrate the isolated security data together. Researcher can build some researching model by some efficient functions of big data technology such as collection, retrieval, storage, analysis etc, and by the related analyses in different aspects and phases, so as to effectively discover the rubbish information, security bugs, ATP attacks and so on, and to strengthen the initiative of network security protection. At the some time, the analysis by means of big data technology can take the best of the huge data resources that is more comprehensive, it can not only utilize the data of the present analysis, but also can get the related data and "virtual" data that are occurred during the analysis, so as to achieve better integration of the data for network security analysis, at last, to make a scientific and effective analysis.

Security big data analysis means the analysis that based on the abnormal situation of the network, by analyzing and modeling the huge data, it can find out the abnormal data and the characteristics of the abnormal data, and it can also design targeted related analysis method for different security situation. The advantage of the security big data analysis is unique, this analysis method can further explore and sorting out the required data, which includes some parts such as data collection, storage,

  DOI: 10.25236/systca.18.026

check and intelligent analysis. Form the aspect of collection, storage and check of the data, the advantage of the big data technology can fully improve the efficiency of network security analysis. By means of big data technology, the researcher can not only collect different kinds of data, but also can collect different data by different ways; for example, Fluma and other tools are adopted for log information, data image method is adopted according to the data size and so on. On this basis, the platform of big data security analysis can realize the analysis from top to bottom, including the aspects of collection, storage, exploration, visualization and so on; according to the data, intelligence information and other multivariate data, it can carry on the distributed analysis, build different models for different data scenes, and strengthen the information management and control of the analysts; these diverse analysis methods greatly improve the efficiency of network security analysis.

2. Specific application of big data technology in network security analysis

From the specific content of the current network security analysis, data and log are the main aspects of the analysis, while data assets, configuration, user behavior, application and other aspects are the related auxiliary information, so we can say, The core of network security analysis is "data"; and the concrete application of big data technology in network security analysis is to use big data to carry out data analysis. According to the steps and contents, it mainly includes the following two points:

As mentioned above, big data technology can process the data in network security analysis by virtue of its unique functions, mainly as follows:

(1) Data gathering

In the gathering of information, we can use the tools of big data technology, such as Chukwa, and use the distributed gathering means of big data technology tools to collect the information that we urgent need; these tools are very convenient and efficient. At the same time, we also can gather the different information by different means, so as to improve the efficiency of data gathering.

(2) Data storage

Different types of data can be applied in different methods, and the different data should be restored in accordance with different requirements. So, in order to meet the requirements of different types of network security analysis and improve the speed and quality of the analysis, it is necessary to storage different data in different ways, so as to achieve more complete ways of data storage. For different data types, such as log information in raw security data, data size, and so on, we can use column storage methods such as Hbase, these methods have the function of fast indexing, and they can retrieve the corresponding data faster. For the data that need to be analyzed immediately, we can use the streaming storage methods such as Storm, and so on, these methods can be applied to calculate and process the data better. For example, the data can be distributed on each node of the storage so as to realize the automatic calculation, and realize the statistics and warning function of the data.

(3) Data indexing

In the traditional network security analysis, because of the huge amount of data, it is difficult to classify accurately, so there are many inconveniences in the indexing of the information, but by means of big data, the efficiency of indexing can be effectively improved, each node in the data query request can be processed according to the requirements, and then the distributed parallel computing method can be used to greatly improve the indexing efficiency of the data; at the same time, it can make the indexing results more accurate, and the results have more analytical significance.

(1) Data feature analysis

In network security analysis, the features of data can be divided into real-time data and non-real-time data, so the types of data feature analysis are real-time data analysis and non-real-time data analysis. In traditional analysis, there might be some problems such as the results are inaccurate or the analysis methods are complex no matter which kind of data feature is analyzed; there problems has a serious impact on the quality and efficiency of data analysis. While by big data

technology can solve this problem well. In the analysis of non-real-time data, we can use Hadoop Structure, HDFS distributed storage and corresponding distributed computing, and then combines data aggregation, data mining, data extraction and other technologies to analyze the attack source of the network from multiple angles and find the position of the attack source at the utmost. In the analysis of real time-data, we can use the streaming computing structures such as Storm, etc. we can combine the complex event processing technology and corresponding algorithm, and then to analyze the real time data immediately, and capture of abnormal of the data more immediately.

(2) Multi-class data association analysis

There are different types of data in the network security analysis, and different types of data have different characteristics. Big data technology can effectively improve the storage performance and analysis efficiency in the process of analyzing different data, and find the data with different characteristics in a shorter time, and find a variety of security risks in different data. For example, in the analysis of the botnet, combined with data, we can not only find access features, but also can further filter and sort out the data by big data technology, finally get more helpful information for analyzing, and then realize the association analysis of several kinds of data. For example, if a host is attacked, the attacked host can be used as a base for checking if other hosts are also attacked, so as to make preparations in advance.

3. Structure of network security analysis platform of big data technology

The platform is the basis of network security analysis by means of big data technology. Under the background of big data technology, the structure of the platform is mainly divided into the following requirements: the first element, data presentation level. Data presentation level is to achieve the processing of raw data by safety measurement technology and other technologies. The most important role of this level is that the raw data can be safely analyzed and processed; the second element, data mining and analysis level. The level of data mining and analysis is the induction of data categories, such as log data, data flow, etc. Because different data have different characteristics, this level can lay a solid foundation for the following analysis work by summarizing the data; the third element, storage level. Storage level is to store different types of data after classification, so as to ensure the smooth development of subsequent work; the fourth element, data collection level. Data collection level is to combine different data to collect the data. Compared with the traditional methods of collection, collection of big data technology is more convenient and efficient.

With the support of big data technology, the implementation of network security analysis is mainly reflected in two aspects:

First, detection of DDOS attack path. The big data technology can detect the attack path of DDOS, and then trace and monitor the attack path, so as to solve more problems in the development. First of all, establishing the empty node, empty attack linked list; secondly, after discovering attack alarm, taking it as the reference object, and accurately extracting the ID, flow, router, etc of the attack alarm object, and use the extracted information to determine the path and carry out security countermeasures. Finally, the router of the attacking object should be taken as the foundation, and the router must be checked carefully, and then to expand the checking scope wider, to find out all the adjacent nodes, and to solve all the security risks.

Second, detect and control. If a host is invaded, by the big data platform, we can detect that the process is: firstly, collecting the history information of the DDOS attack, identifying the IP of the attack, and classifying the attacked host as "suspected". After listing the "suspected", the log of the invaded host should be detected, and the detected result should be compared with the URL. If there are similar intruded logs in URL, then the source of attack can be checked out. On this basis, in the process of being attacked, we can also use the technology of multivariate data association analysis to find the IP information involved in the attack in a single attack period, to realize the detection of association, and finally to find the corresponding solution, and finish the network security analysis.

4. Summary

Generally speaking, because of its unique performance, big data technology has become more important in network security analysis. By the influence of the characteristics of Internet, it is very difficult to analyze and check the security of network at present. Only by means of big data technology can the efficiency and quality of network security analysis be improved more effectively, and can the development of network security in China be realized.

References

[1] Wang Shuai, Wang Laifu, Jin Huamin, etc. On the Application of the Big Data Technology in the Network Security Analysis [J]. Telecommunications Science, 2015(01).

[2] Jia Sunyu. Brief Discussion on the Application of the Big Data Technology in the Network Security Analysis [J]. Network Security Technology & Application, 2017(11).

[3] Cui Yuli, Huang Lijun. On the Application of the Big Data Technology in the Network Security Analysis [J]. Cyberspace Security, 2016(04).