Research on the security technology of big data information

Hong Zhu[1], Zheng Xu[2] and Yingzhen Huang

The Third Research Institute of Ministry of Public Security

Shanghai, 339Besheng RD, The China

[1]zhzj339@163.com, [2]dolly_xz@hotmail.com

Keywords: Information security, Big data, Network attacks, Stealth viruses, Intelligent network management.

Abstract. The arrival of the era of big data, which is very important for building a good big data security system and improve the system of extreme defense measures. This paper presents six major issues of information security in the era of big data. It introduces the security technology of big data system, including the security redundancy of the hardware system, the improvement of the traditional information security technology, the active analysis of large data mining stealth virus, and Intelligent network management to clean up the network environment. It discusses the stealth virus is modeling, mining, garrotte technology, and the future development of information security technology are presented.

Is the system of the age of big data wealth safe?

In today's world, human activities and all aspects of the whole society are actively or passively through the data in the record, covering a variety of information sensor to the social movement of the data to mark or save. Has formed a world of information and information to live. Every day in all kinds of systems, such as a large amount of information data, so that all walks of life are starting to build big data applications platform. To reveal the information rules which contain large amounts of data within the industry, mining large data contains a huge amount of wealth from. The new type of big data platform can monitor and forecast the dynamic and changing nodes, so as to control and adjust its operating rules. Create a variety of visual charts to reveal the inherent relationship of various data in the industry, and promote business value. Wisdom statistics, decision, give advice predicted events that may occur in the future or to predict when and where will ensue and intellectual property management. Therefore, now the information industry has announced a major event: the era of big data wealth.

Big data platform for large data processing capabilities in the world to surprise the world, but also in the huge security risk of large data presented in front of you. International concern: large distributed data platform software in multiple attacks against hackers will numb and vulnerable? Are all kinds of network security isolation facilities become vulnerable because of the complexity of the network? Large scale big data hardware facilities need to address the security and redundancy of the data backup of these requirements, the amount of investment will be limited to live? To the big data era personal privacy can be protected? Because the big data platform has been a business or an industry's key data information in the system. In this era is the data is wealth, if you can't protect large data system, the occurrence of a collapse or the outside world to capture and other tragedies, which will be big data has a big risk of the terrible consequences. Therefore, the era of big data security data security is an unprecedented huge, social security needs of large data system [1]. Building a good big data security system and improving the system of good system of extreme defense measures will be extremely important.

Equipment safety of big data system

Big data information security must first system equipment security. Most of the big data system is built on the Hadoop, Sparke and other distributed computing platform, the data collection, cleaning, service, storage, analysis, reveal the rules and intelligent forecasting. The scale is very big, and it is not limited to added the server running on the system bus. The distributed parallel operation has a big number of controller and memory hard disk memory, system protocol and network material and quality are very important. When the big data system to reach thousands of servers, the use of Service Web and other services architecture is very complex background system. There should be a safe and redundant backup of the device, to ensure that the calculation of a big number of data, the calculation of the value of the interface and the results give the user a variety of interfaces. Big memory is the place of the distributed high speed calculation and cache database. It is very important for the system to run safely. A number of hard to build the data warehouse data, for massive data storage and retrieval, should use high-end long-term hard disk storage. Use mechanical drive hard drive must be fully backed up and intelligent inspection. Data about the core information, the establishment of a reliable big data storage warehouse, although it is also necessary to. In the development of high memory, big data control manager should pay attention to the problem of radiation interference and delay error when running the big amount of data. Provide remote visual management and operation of the interface and components, improve the operating environment and no fault time, the construction of a reliable and effective system for big data hardware security[2].

Big data platform must have a reliable and secure redundancy system architecture. Through a variety of controlled routers and firewalls, data switches and other facilities to establish a clear boundary and data interface management; data encryption protocol, different data structure to have the corresponding conversion processing module, the establishment of backup security redundancy. The hardware equipment of big data system should improve the reliability and anti-interference ability, through the research and development of trusted chip, trusted bottom board, trusted memory and trusted server, eliminate the leakage of hardware and improve the processing capability and security. In the face of the most vulnerable to the external attack of the channel interface, big data system must be in the prevention of light electromagnetic information leakage and other aspects of a more effective response to defense measures. These are subject to big data system construction to the extent of the degree of safety and construction funds and the adequacy of the dual decision.

The traditional information security technology about log, level protection assessment, etc., can be well applied in the large data system of information security and operation and maintenance. Big data system are from the collection of data cleaning, sorting, clustering search to form clusters of data, and pattern recognition, statistical computing to implement mining analysis of judgments. Finally, the visualization analysis results or data information law. The whole process for the size of the hardware system may be very large, often requires a lot of equipment in a large number of cabinets, power consumption, high environmental temperature requirements. The mutual interference between each device can not be ignored. The quality of each device is related to the operation of the system. Therefore, the operation and maintenance of large data systems is complex, large scale, technical requirements higher. Should give full play to the advantages of large data processing and information mining, in the system to increase the operation and maintenance of data collection and operation of intelligent data accumulation. Automatic analysis of the running trend of hardware equipment and the operation of the system, the establishment of a variety of fault models, intelligent monitoring system operation and early warning of the risk of failure to achieve active intelligent operation and maintenance.

Big data information security environment and response

Information security of big data is a long-term continuous development of technology, in the future, the advent of a revolutionary technology to change. Now big data applications are faced with the following security: 1, the scale of the network to attack the huge damage, this system, with the top integrated technology to support the attack means to protect the network and the platform has become vulnerable; 2, a variety of different access devices and distributed database and handheld mobile devices such as the huge amount of nodes of large data systems to form a large and more vulnerable network; substantial changes in 3, big data scenes of the separated trojan virus tends to be invisible, resulting in real-time security monitoring and defense difficult subject; 4, the data system is often composed of many different security levels of the network, the inter network isolation and data flow by directional performance related equipment limitations; 5, the wealth of data will attract countless hackers attention, and the number of intrusion system according to the leakage means it is possible to go beyond the scope of the network now, system physical isolation will by way of light, electricity, magnetic violations and attacks; 6, big data can reveal personal wealth, personality preferences, circle of friends and many other privacy, personal privacy is the need for legislation to protect.

These information security issues is a big challenge to the era of big data wealth. On the current big data system construction, cloud computing strategy, a new integrated network, the wisdom of the land can not be underestimated. Will be directly related to the large data information security development direction. Therefore, the work of big data information security, is one aspect to improve the traditional information security technology defense. To establish a more stable boundary and channel. On the other hand, it will study the system analysis of large data and machine learning application in information security. Risk sources through network sensing technology. Initiative to destroy the hacker program, the initiative to clean up garbage, intelligent network management.

Comprehensively improve the traditional information security technology. Through the development of new hardware to protect, improve the access and control of the system and the key management, authentication and trusted channel and other protective measures. Strictly regulate the use of data and communication protocols for distributed systems. Strengthen the traditional techniques of interlocking mechanism and update sequence, convergence state, log, security time, redundancy recovery and so on. Makes information security level protection mechanism and technology to be fully improved.

The information security of large data needs a multi-level security policy model and prevention system. Conduct multilateral security division, model reasoning and physical protection. Monitor network attacks and network protocol vulnerabilities. Fine analysis of the source data for modeling and mining. Screening to combat and prevent the emergence of the intrusion and the formation of stealth virus, virus and multi - level calculation of the virus. Global defense of all kinds of network attacks. Control and safe operation through configuration management, encryption, topology, optical gate, firewall, etc.. By filtering the data flow, intrusion detection, privacy protection technology, system security and safety assessment. To establish a reliable protocol coordinator and participants of the distributed system and database and log. Strengthen the security of log management at all levels, border and channel. Take the one-way path of light brake to resist invasion. Also pay attention to the threat from all kinds of possible. Often find the weak link of the system, to take a variety of technologies, multi pronged, closed [3].

This paper puts forward the information security measures of large data system: 1,strengthen the design and demonstration of the system structure of the large data system, ensure that the system has the ability to resist all kinds of external attack information security measures, and the use of trusted devices of large data platform system. 2, in different security levels of the network to control the flow of data between the border and the nodes of the data analysis of data analysis of the security of the tunnel, the use of reliable one-way data flow control to ensure that the low level of network security can not attack the high level network, found that the problem can be smart enough to pursue and attack the exclusion, to ensure that the source of data is reliable. 3, to control network source, the establishment of a sound system of large data security data security scanning mechanism, the use of a variety of data mining information security attack model, cut off the threat of the sub body and stealth virus attacks, intelligent cleaning garbage and purification network. 4, strengthen order of big data network, in the vast network of established inside and outside of multilevel network management and multilayer intelligent anti-virus barrier, using data analysis and forecasting methods in the network system take the initiative to find and strangle the virus, can take the initiative to attack the suspicious program on the modified so that the loss of damage. 5, to strengthen the data flow between different security levels in the large data network data flow encryption measures to protect the wealth of data and other important data is not recognized. On the high level of network and physical isolation of the network, to take the photoelectric magnetic field shielding technology, to prevent the leakage of information from the back door of the device. 6, the legislation of privacy protection measures, combined with network real name system, network identity and data linked against various privacy stealing and leaking.

Exploration of active information security technology

In large data information security system, the data acquisition, fusion classification, judged the modeling, analysis of early warning and intelligent global strangulation to take the initiative to attack viruses and network attacks. Network attack in addition to the forming of the virus, Trojan horse, there are many seemingly normal data or word set pieces, and approximation perturbation of some interference sources, can cause damage to the system. To establish a detection model for these threats, the collision of the data stream at each node and the system entrance. Now, the visualization of large data analysis and prediction technology can dig out the event rule in the data, which can predict the development trend in the embryonic stage. Using active information security technology, we can find out whether the data is associated with the virus in the early time, and tap the source of danger. Through intelligent network management initiative to clean up the garbage and eliminate virus hazards, to prevent loopholes and intelligent[4].

Building large data active information security system. Through the research and development of large data intelligent network management, scanning analysis system of garbage, mining and analysis of the various forms of hidden traces of viruses and their formation. Monitor and dig out the signs of the virus attack in the data stream. Proactively and timely response to garrott. The network security environment of large data is cleared by the intelligent cleaning of the garbage. Maintain the normal work of the hardware system, in particular, to protect the security of the multi core controller and the work of the distributed large memory system. Maintain system space and improve operating speed. Maintain the best efficiency of the system, so as to achieve the system's information security, active defense and eliminate all kinds of intrusion.

Summary

Along with the value of the big data wealth is more and more, people's demand for the information security of big data will be more and more high. Make full use of the advantages of big data technology to information security. Analysis and forecast of the means and development trend of various system attacks. The system of killing the virus and the threat will be more intelligent and efficient through machine learning. Large data information security technology will become increasingly important.

References

[1]Information assurance technical framework, Rclcasc3.1[S].2002.

[2]ISO/TEC15408-1999,Common criteria for information technology sccurity cvaluation V2.3[S].2005

[3]Viktor Mayer-Schonberger,Kenneth Cukier.BigData:A Revolution that Will Transform How We Live ,Work and Think . Boston :Houghton Mifflin Harcourt,2013

[4]Carr,J.(2010).Inside Cyber Warfare,O Reill and Associates, Inc.Sebastopol,CA.