Discussion on Information Security Technology of Big Data System

Hong Zhu, Ju Wu, Huan Shi, Changning Wang, Yahui Li
The Third Research Institute of Ministry of Public Security, Shanghai 201204, China
zhzhj339@163.com

Abstract. This paper points out the importance of information security in big data systems and the importance of improving the equipment and equipment production standards, and builds the system architecture of trusted equipment system and security redundancy. The method of eliminating vulnerability and resisting virus threat of big data system software is discussed. Study on the prediction and analysis of intelligent network management system for security and safety inspection of pattern recognition and fault situation, with big data intelligent security technology to run the program and data on the internal flow path and the boundary interface channel for safety monitoring, virus attack source model analysis, mining and clean up the system to strengthen the anti-virus, data encryption and network identity research detection. Through the early warning technology of information security technology and mining analysis to analyze suspicious intrusion early warning, data acquisition, discovery, and all kinds of threats and attacks to avoid anputation of system risk and virus source, to achieve big data high efficiency range of information security system.

1. Introduction

In today's world, human activities and all aspects of society are recorded actively or passively through data. Information sensing equipment covers all kinds of technology to human and social activities of production and management of massive data collection and analysis, and using the machine learning depth novel artificial intelligence technology for data mining connotation again. For the embryonic world of intelligent wisdom, the information world and data that have relied on information have become the new trend of social activities. Each kind of system produces each day rich and colorful each kind of information, the data has become the resources, has caused the all walks of life to establish the big data application platform.Is using a variety of data within the industry and big data platform through analysis of deep-seated rules hidden behind big data industry comprehensive summary, to reveal the key industry operation of the multidimensional information and predict growth and industry fluctuations weak position. So as to excavate the great wealth contained in the big data of this era. This has led many companies to make money by investing a lot of money in creating new platforms for social data. The key to making money in disguised form is that data belongs to one kind of wealth. Big data platform with super value-added capabilities, monitoring, forecasting industry trends and changes in the operation of nodes. Master the operation rules and the relationship, promote the business value, wisdom, statistics, decision-making advice, predict possible future events or control when and where and what happens to achieve wealth management.

But big data systems also have huge security risks. Big data platform income of a variety of industry or social critical data, industry operations and personal privacy in the big data systems and networks. Due to the complex huge system and network, many hackers in this world, full of threats,
the global Internet network will enable each big data system faced by all kinds of threats and attacks from internal and external. Science and technology are developing and new technologies emerge in endlessly. New intelligence and stealth attacks, multimedia fishing technology, high-speed computing and crack technology have emerged, which have caused the escalation and fierce attack of hacker attacks and viruses. If there is no effective large data system, information security technology, hardware devices can not be trusted, the network, software systems and interface protocols have a variety of loopholes. So that big data systems can not be safe and reliable protection of big data, information and system security, system crashes, leaks, or captured by the outside world and other tragedies, the loss and disaster of the general will be a terrible [1]. Therefore, we urgently need to strengthen the protection of the situation has a big data security system, improve information security measures, killing the virus invasion, resisting risk and research system for the future quantum computer not far from the spread of failure impact remedial measures of the traditional encryption technology.

2. The key to large data systems is to design and plan information security
Big data system brings together all kinds of data for analysis and processing, which brings people the hint of the information results and inherent laws. Summarize and clarify the clutter data. Form valuable information at all levels. The big data system contains the core of the data involved. In this chaotic environment of global integration, information security has become one of the keys to big data systems.

The information security of large data systems is faced with internal and external threats. Internal defects are mainly caused by software vulnerabilities, program BUG, system design errors, and poor hardware performance. This is because developers or system designers, software and hardware vendors and systems integrals have low technology levels or insufficient development time. It is also the reason why the responsibility of the supervisor is not in place. It may also be the result of the negotiation failure or lack of investment by the user when determining the system goal. But mainly the system design and development of technical level is not enough specific performance. It needs to be resolved through further system improvements, software improvements and increased investment. Avoid greater losses. External threats are threats from external viruses, hacker attacks, and running garbage conflicts that cause information to leak or crash. General system garbage can be automatically cleared by management software. But external attacks, if not defensive, will cause serious damage to big data platforms and systems, and huge losses, [2].

From planning and designing large data systems, we should pay attention to and do a good job of the entire system of information security, clear responsibilities. In the architecture design, the information security of the basic layer and the data layer can be well protected, and the data and the processing process at all levels can be protected from infringement. In the service layer and application layer to do the interface management and password protocol settings, to take physical methods to cut off or block outside attacks and intrusions. Support layer to do a good job of viruses and various threats scanning and removal, through the new big data intelligent network management system for the protection and early warning maintenance. For large data platform, system network, data acquisition points, data warehouse, output display terminals and other parts of the system, software and hardware system information security protection should be carried out. Based on the traditional information security protection measures on the need in accordance with the complex system of big data network information security system, intelligent protection and take the initiative to clean up the external threat kill series, big data and information security, to achieve the purpose of active information security model.

3. The construction of large data systems to use credible and reliable hardware
With the advantages of distributed large system technology, most big data systems are now composed of Hadoop and Spark distributed data platform. It can connect thousands of servers on its system bus and mount numerous terminals to run in common or distributed. Numerous nodes are connected to
many databases, data acquisition or terminals, and can be made into large scale data systems and networks. Such large-scale data hardware system, often requires a large number of large cabinets of equipment composed of data centers, power consumption is great, the environment temperature and power supply requirements are high. A large number of servers include backplane, controllers, memory, power supplies, fans, connectors, and many other devices and systems. Because of the complexity and complexity of the system operation, the system is of high standard, high quality and applicability. The equipment should have better anti leakage, anti electromagnetic shock, power fluctuation and temperature adaptability. These require reliable system architecture and reliable equipment and equipment support, and security redundancy of equipment backup and comprehensive operation and maintenance system to provide protection. In order to ensure the long-term work of big data systems and to cope with the adverse environment of the ability to resist risks. Reduce the probability of failures and downtime. Achieve the functions required for the construction of this system. Complete data acquisition, operation, analysis, data cache, result display, data storage and other data flow. At all levels of instruction, according to the requirements of the uninterrupted service of the various nodes and interfaces[3].

System equipment is safe and reliable, is the foundation of information security of big data platform. Benefit from the rapid development of related basic industries. Establish big data system equipment safety standards, according to this high standard implementation hardware equipment production. Greatly improve the reliability and reliability of the system hardware. Industry to big data analysis, hardware security, performance, quality and reliability testing. As technology upgrades, more stringent standards for hardware are established. The introduction of trusted high-performance servers and high level network devices, as far as possible to eliminate the hardware radiation leakage. Improve the performance and safety of equipment. Establish a comprehensive, safe and reliable new testing system for related equipment, equipment and whole machine. Make domestic and imported related equipment have more safe and reliable testing standards and the basis of use. Vigorously study and develop new large data security equipment. Promote high security router, one-way all optical gate, intelligent security network management equipment replacement. China is in the leading position in quantum technology. We should speed up the study of quantum cryptography security devices, and strive for quantum security encryption era as early as possible.

In big data system platform design, planning and integration. Through the use of high security level routers, optical gates, network systems, intelligent security management and control equipment, quantum encryption network equipment. The system can protect big data systems through one-way data channels, more rigorous physical isolation, system security, intelligent management and new encryption systems. To be free from disclosure or invasion. This is through the reliable hardware technology carries on the information data security protection. The research and development of various system boundaries, monitoring channels, interface nodes, data encryption key protection, security, credible, non leakage of various chip equipment and complete measures, broadband security network structure. The implementation of the key data of complex network intelligent protection, clean up the garbage, virus, strike against external attacks. Raise the system's level of protection against new threats. Improve and strengthen the key combination of routers and new firewalls. Active monitoring and troubleshooting data network channel devices. The isolation measures of deepening physical isolation and electromagnetic information leakage make the traditional information security technology updated and improved.

4. Large data systems software security technology upgrades

Big data systems should upgrade system software for information security technology. Strengthen the reliable design and demonstration of the big data system architecture to ensure that the system is able to withstand various attacks. In different security classes, there is a secure data tunnel between boundary isolation and node data analysis to control the flow of data. Both static and dynamic data should be encrypted, the key separated from the encrypted data system, and the use of multiple network identification measures and protocols for strict control of the interface and data stream. A
network with a low security level can not attack a high level network and ensure data is reliable and not compromised. Data tunneling modeling and collision analysis of data acquisition. The discovery of problems can intelligently eliminate and trace the source. Establish a complete system information security scanning analysis mechanism. By the analysis of big data mining prediction techniques to deal with the split virus and stealth attack, early discovery and active kill virus. Ability to proactively attack a variety of suspicious programs and modify them to render them less disruptive. Strengthen the order of big data network, establish internal and external multi-level network management and multi-layer intelligent anti-virus and antivirus barrier in huge network. The system cluster has automatic security settings and event logging. With new intelligent network management and operation technology, intelligent elimination of various threats, garbage removal and purification network. Secure communication is used between nodes and between nodes and applications. Strengthen the protection of important data information in big data network. For high level networks and physically isolated networks, photoelectric field shielding technology is adopted. Prevent all kinds of back door leaks. Combined with network real name system, network identity and data hook. Combat all kinds of privacy, theft and disclosure.

The architecture of big data software systems is more complex, large, and more hierarchical. Multiple heterogeneous data aggregation to platform processing. To solve cross platform data access and sharing, the internal management of data transmission module, data encryption and decryption module and data correctness check module are carried out to ensure the normal security of system processing and data exchange. In distributed systems, middleware is used to automatically identify data types and security. Implement data encryption and decryption of each system transmission, automatic navigation and transmission. By shielding the differences between operating systems and network protocols, the application provides a variety of communication mechanisms that enable large data platforms to meet the needs of different domains. Large data systems need to be transformed from outdated security policies and protocols. According to the big data, complex distributed system structure, a new model of virus and hacker attack, to improve the level of information security protection standards. The data stream of the data channel carries out the key protection of access authentication management and data encryption. Perfect traditional information security technology. Enhance the ability of new firewall to resist external attack. In a complex network of intelligent monitoring, fine freezing, active strangulation and automatically clean up internal network and external virus attack. Monitoring the presence of data penetration between physically isolated networks may create more stable boundaries and channels. Effectively overcome the shortcomings of traditional information security.

Large data systems, information security, network management software, system intelligence scanning, monitoring data flow to carry the external threat of attack data stream and useless garbage, stealth attacks and virus deformation attacks. Intelligent and active clearance, effective management of the system. As a large scale, complex format, rapid processing and complex large data platform, it is more vulnerable to external targeted attacks. Various virus attack models need to be developed to identify and analyze them. Research on stealth virus identification and one-way data channels. To take the initiative to identify and remove the artificial intelligence method, blocking Trojans and hacker attacks. Perfect border isolation to prevent system errors and data leakage. Applications for distributed system software on large data platforms are connected to multiple networks of different security levels. The core network needs to be secure and reliable through data interaction monitoring between the effective boundary and the security interface, screening trusted data and strictly isolating the virus. Large data interface scenarios are extremely complex. Kill viruses, viruses, Trojans, and stealth attacks by killing viruses and viruses. To achieve fast intelligent perception and system security monitoring, effective anti-virus and network management [4].

5. New information security technology research
With the mobile phone has become a comprehensive human-computer interaction media. The popular Internet of things makes things around you intelligent and controllable. Promote big data, mobile
Communications become the protagonist of network development. Coupled with the future can not be cracked by the quantum encryption network, 5G communications network to join, various ways of satellite communications, mobile Internet and other new technologies to join. Use security software to protect mobile phones from infringement. The application of quantum security mobile phones and quantum wireless networks in the future of quantum key protection will upgrade the security of mobile networks.

The new type of big data security system functions as a system for security inspection, antivirus and threat analysis and early warning. Improve the system's data access check and access handshake key agreement. More difficult to decipher the data encryption and key management, the establishment of trusted channel. The model inference and physical protection of multilateral security segmentation are carried out by using multilevel security policy model and guard system. Set the level of information corresponding to the level and level of the interface. Corresponding to the strict living fingerprint, face biometrics and other multi entry inspection authentication system, the whole network popularity network identity system. The system program should monitor and analyze all kinds of vulnerabilities and threats comprehensively and intelligently. The use of data and communication protocols is strictly standardized for distributed systems, networks and mobile systems, and in a variety of scenarios. Internal and external fine monitoring and searching, application security, interlocking mechanism and update order, convergence status, network security. Fine analysis of path and node of the source port data mining modeling, through the intelligent analysis system to coordinate inventory of all kinds of attacks and exterminating the virus. Control through service configuration and operation security management, encryption, topology, firewall review. Trusted security algorithms and computing processes, data flow monitoring, and finding weaknesses in the system. System guarantee and security assessment, multi pronged approach, closed loopholes. Distributed systems and databases establish reliable protocols for coordinators and participants. Screening, combating and keeping in mind the invasion and generation of stealth viruses, sub viruses, and multi-stage computational composite viruses. Global defense, network attacks of various scale levels [5].

Big data systems to establish intelligent network management and operation and maintenance of big data, large equipment inspection system, log and patrol maintenance. Each device and node of the system should set up the monitoring point. Through the intelligent operation of the system, intelligent log inspection and operation and maintenance mode, the purpose of intelligent operation and maintenance is achieved. Research on equipment operation data monitoring and intelligent operation and maintenance mode, self collection and control solution. Master the equipment operating environment and run the relationship between various fault curves and data. Establish the analysis model of system failure early performance comparison model and fault pattern recognition, relying on cloud service, remote device intelligent inquiry and log supervision. And interface services, protocol, certificate monitoring, etc., constitute a reliable and effective hardware support system for big data systems. Using big data mining technology, automatic processing and analysis of running data of each device are processed. Through the operation and maintenance model algorithm, calculate and judge the situation and time of the occurrence of the fault. According to the operation and maintenance plan and instruction, the intelligence will avoid, notify or automatically activate the redundant equipment and troubleshoot the equipment. Intelligent operation and maintenance system can greatly improve the security and efficiency of big data systems, and reduce the pressure on users, operators and construction parties. As the big data system operation and maintenance of high complexity, large scale, intelligent operation and maintenance of higher technical requirements. The development of intelligent operation and maintenance should give full play to the advantages of big data processing and information mining. In the system, the relation between data of operation and maintenance, data of each acquisition point and fault curve is studied. Data collection that can automatically collect and analyze equipment running status. Intelligent analysis of the running trend of hardware equipment and operating rules of the system. Establish various fault models, intelligent monitoring system operation and early warning of fault risk, and realize active intelligent operation and maintenance.
6. Ending
More and more attention has been paid to big data systems, and the scale and value of big data systems have risen substantially. The threat of attack is also increasing. It puts forward controllable and inviolable requirements for the information security of big data systems. Prevent the system from invasion, destruction and information leakage. The information security of big data system is a technology that has been developing for a long time. In the current battle between intrusion and defense, the quantum security technology will continue to be successfully applied. To mutate into a new generation of information security technologies. Quantum computing technology is growing rapidly at home and abroad, and it brings a new generation of safer quantum networks. Also has the absolute impact to the traditional network security encryption, the traditional password is instantaneous crack before the quantum computer. But in the transition stage, the security threat of traditional network has become a common problem that must be paid attention to and solved.

References
[1] Viktor Mayer-Schonberger, Kenneth Cukier. BigData: A Revolution that Will Transform How We Live, Work and Think. Boston: Houghton Mifflin Harcourt, 2013.
[2] Comprehensive National Cybersecurity Initiative and United States Naval Institute, "Dealing with today's asymmetric threat, Cyber Threats to National Security", CNCI and USNI, 2010.
[3] JOSHI J, GHAFOOR A. Digitgovernment security Infrastructure Design Challenge[J]. IEEE Computer, 2001, 34(2):66-72.
[4] SUN Y L, YANG Y F. Trust Establishment in Distributed Networks: Analysis and Modeling [C]. IEEE International Conference on Communications, 2007: 1266 — 1273.
[5] Zech P. Risk-BasedSecurity Testing in Cloud Computing Environments [C]. Proceedings of the 2011 IEEE 4th International Conference on Software Testing. Berlin, Germany, 2011: 411 — 414.