Research on Key Technologies of Industrial Internet Data Security

Min Yan¹, Haibo Huang²³, Jun Li³*

¹ Trusted Computing and Information Assurance Laboratory, Institute of Software Chinese Academy of Sciences, Beijing, 100190, China
² Department of Economics and Management, Beijing University of Posts and Telecommunications, Beijing, 100876, China
³ China Industrial Control Systems Cyber Emergency Response Team, Beijing, 100040, China

*Corresponding author’s e-mail: yanmin20@iscas.ac.cn

Abstract. At present, industrial Internet data is facing a severe security situation in storage, transmission, and use. This paper proposes data security management and security protection strategies by studying the data security risks and status quo of industrial Internet platforms. Among them, the security management strategy includes data classification and classification, data security assessment and data security threat emergency management. The security protection strategy deploys industrial data security protection from the technical aspects of collection, storage, transmission, processing, sharing, and deletion, providing a reference for further strengthening the industrial Internet data security.

1. Introduction

Industrial Internet data is an important resource for the development of industry. Data security protection is related to the development of the industry, especially part of the data is the operational data of the country’s critical infrastructure. The security, confidentiality and availability of these data have been destroyed. Not only will it have an impact on the production and operation of the industry, but it may also have serious consequences for the economic development, the ecological environment, and the development of public welfare. However, the current situation of data protection in many industries is worrying. The network security protection capabilities of Internet platforms are weak, and many important data are exposed in the system. It is urgent to strengthen the data security protection of industrial Internet platforms. This paper will elaborate on the status quo of industrial Internet data security, security threats, security management and security protection technology.

2. Definition & Characteristics

Industrial Internet data refers to the typical intelligent manufacturing model in the industrial field, from customer needs to sales, orders, planning, research and development, design, process, etc., the general term for all kinds of data and related technologies and applications generated by all links in the entire product life cycle[1]. With product data as the core, it greatly extends the scope of traditional industrial data. In addition to the characteristics of traditional big data, industrial Internet data has the following six typical characteristics, such as Polymorphism, Real-time, Reliability, Closed-loop, Cascading and High-value.
At present, industrial Internet data has stronger professionalism, relevance, process, timing and analysis, focusing on the physical meaning behind the features and the mechanism logic of the relevance between features, while traditional big data tends to rely solely on statistical tools mine the correlation between attributes. Industrial Internet data has the following differences from traditional big data in terms of data life cycle links and applications. See Table 1 for details.

| Phase             | Traditional Big Data                              | Industrial Internet Data                          |
|-------------------|---------------------------------------------------|---------------------------------------------------|
| Data collection   | low timeliness                                    | real-time collection                              |
| Data Storage      | Weak correlation & free storage                    | highly correlated & complicated storage            |
| Process           | irrelevant data is removed.                       | Data process requires the authenticity, integrity and reliability |
| Analysis          | Coarse-grained analysis result                    | Complicated data modeling, special algorithm and reliable analysis results |
| Visualization     | Generalized Visualization                         | real-time warning and trend visualization.        |
| Closed loop feedback | Generally, closed-loop feedback is not required. | Emphasize the closed-loop nature to realize process adjustment and automatic control. |

3. Current status & Risk
With the continuous development and maturity of new technologies and applications such as cloud computing, Internet of Things, and big data technologies, the deep integration of industrialization and informatization has improved the production efficiency and management efficiency of industrial enterprises, and created more business models. There are many new network security and data security issues.

3.1. Standardization status
China has issued a number of national and industry standards for industrial control security, and there are many industrial control data security standards under development. The standards under development basically cover the standard requirements of industry regulatory units, industrial control product suppliers, industrial control safety protection product suppliers, industrial control owners, and evaluation agencies. In order to achieve real standards for industrial control data security, it is necessary to accelerate the development of standards and release them as soon as possible. In June 2018, the National Standardization Administration of China issued (GB/T 36466-2018) "Information Security Technology Industrial Control System Risk Assessment Implementation Guidelines". On August 30, 2019, the National Standardization Administration of China issued (GB/T 37962-2019) "General Information Security Technology and Industrial Control System Product Information Security Evaluation Guidelines." [2].

3.2. Technology risk
Industrial Internet data includes all kinds of data in the entire product life cycle from customer needs to sales, orders, planning, R&D, design, technology, manufacturing, etc., and there are many security risks in the entire life cycle of collection, storage, transmission, analysis, and application as follows:

3.2.1. Data collection security risks. Traditional PLC, CNC machine tools and other systems and equipment may not have the data upload function, and data acquisition gateways and sensors are installed. Waiting for data collection equipment, or converting data interfaces, may affect the performance of the system and equipment, and there may even be a security risk of data leakage or tampering[3]. At the same time, the lack of data collection interfaces and protocol standards makes it
difficult to uniformly deploy security protection measures. In addition, once the collected data carries security risks such as viruses and Trojan horses, it will also affect the industrial Internet data security.

3.2.2. Data storage security risks. In the process of data storage, the effectiveness of security protection measures such as boundary protection, access control, and regional isolation is reduced. The distributed open cloud architecture for storing industrial Internet data leads to blurring of system boundaries, and traditional security protection measures based on boundary protection are at a discount; because industrial Internet data application scenarios such as electric power, petroleum and petrochemicals, and equipment manufacturing require specific permissions for a large number of users, it is difficult to accurately specify the scope of data access for each user; the distributed data resource pool in the industrial field gathers a large number of industrial data. These data are large in scale and diverse, and it is difficult to implement regional isolation measures.

3.2.3. Data transmission security risks. In the past, the data transmission protocol between industrial systems and equipment was closed, security considerations were insufficient, and encryption measures were lacking. Taking into account the real-time and polymorphic requirements of industrial Internet data, traditional high-strength encryption, homomorphic encryption and other measures are difficult to apply, and there is an urgent need to develop efficient and lightweight new encryption technology methods. Under the trend of industrial interconnection, data transmission between distributed industrial network data nodes and between industrial Internet data-related components is facing security threats such as monitoring, theft, tampering, and large-scale leakage.

3.2.4. Data analysis security risks. Industrial systems and equipment such as industrial production equipment, sensors, and smart terminals continue to generate a large amount of industrial data. The source data is multi-dimensional, heterogeneous and fragmented, which enables traditional security protection such as data hierarchical storage, cleaning and analysis, authenticity verification, and low-disturbance analysis. The implementation of measures has become more difficult. However, industrial Internet data has high requirements for refined analysis and extremely low fault tolerance, and it is urgent to develop targeted new security protection technologies and tools.

3.2.5. Data application security risks. Big data, cloud computing, industrial control, artificial intelligence, machine learning, and other interdisciplinary fields of technology integration applications face high technical and management risks; industrial Internet data applications mostly use third-party open source components, these components lack rigorous testing Management and security certification make industrial Internet data applications insufficient to prevent software vulnerabilities and malicious backdoors.

3.3. Managing risk

3.3.1. Weak awareness
The safety consciousness of industrial enterprises is not enough. A large number of enterprises do not know how to carry out security protection work, nor how to evaluate the effect of security protection, and have not fundamentally solved the industrial Internet data Security issues. There is a security risk that internal personnel use their own access rights to steal or tamper with data. The ability of industrial enterprises to control data and business systems is weakened. In the traditional model, the data and business systems of industrial enterprises are located inside the enterprise, under its direct management and control. In the industrial Internet environment, the security of industrial data and services mainly depends on the platform providers and the security measures they take, and it is difficult for industrial enterprises to understand the implementation of these security measures.
3.3.2. Weak supervision
The responsibility of the security subject of industrial big data is unknown. In the process of sharing, exchange, and circulation of industrial big data, data owners and managers are different, and data ownership and use rights are separated, resulting in data abuse, unclear ownership, and unclear responsibilities of security subjects. Risks will seriously damage the rights and interests of data owners. The ability to guarantee industrial big data security technology needs to be improved. The survey found that industrial big data is still in the stage of promotion and development, security has not kept up, and the "three synchronizations" have not been achieved. The existing security measures are mainly traditional security protection methods, and specialized industrial big data security technical methods relatively lacking.

4. Industrial Internet data security strategy
Industrial internet data security strategy includes two aspects: security management strategy and security protection strategy, as shown in Fig 1.

![Industrial Internet data security strategy diagram](image)

Figure 1. Industrial Internet data security strategy

4.1. Security Management Strategy
Security management strategy includes three aspects: data classification and classification, data security assessment and data security threat emergency management.

4.1.1. Data classification
Data classification aims to sort out the massive data assets of industrial enterprises through classification, clarify the basic data types, determine the sensitivity of various types of industrial data,
clarify the scope and boundaries of the data and use methods, in order to strengthen data security management and promote open data sharing And maximize the excavation and utilization to provide support. Enterprises should classify and classify data according to the importance of data, and establish and improve data information management catalogs and supporting management systems on this basis. Encrypted storage or isolation protection for important industrial data, set up access control functions to ensure that important data is not illegally accessed, deleted or modified, and keep approval records for legitimate access processes.

4.1.2. Data security assessment
Data security assessment should formulate inspection and assessment system specifications, establish a data security inspection and assessment team, and regularly carry out data security inspections, assessments, random inspections, and supervise rectification in accordance with procedures; regular security inspections and comprehensive inspections of important and restricted data should be carried out regularly Inspect the evaluation report, and report the results of the inspection and evaluation; according to the inspection and evaluation situation, take timely measures to make rectification and review.

4.1.3. Emergency management
Emergency response relies on emergency resource data to realize the comprehensive management and compilation, generation, and transmission functions of business instructions and control instructions, support emergency command communication, and realize video conferences, including one-to-one, one-to-many, or many-to-many. On-site emergency command, instant messaging and other functions, to achieve the ability to record the entire process of command, form event reports including perception, processing, and evaluation of the entire life cycle, integrate data from various systems and modules in the emergency response platform, and highlight resource allocation and developments , The operation status of the command channel, the implementation of response measures and the effect evaluation, and grasp the emergency command actions throughout the life cycle.

4.2. security protection strategy.
Focusing on the data security protection requirements of industrial enterprises involved in customer needs, design, R&D, manufacturing, sales, service, recycling, etc., industrial enterprise data security protection is deployed from the technical aspects of collection, storage, transmission, processing, sharing, and deletion.

4.2.1. Data collection and border protection
According to data classification, the network where different data domains are located should be divided into different network areas, and boundary protection measures should be established. Data interaction and access across the border should be communicated through the controlled interface provided by the border device. Data collection security control include:

1) Take necessary testing, certification and other measures for the data collection environment, tools, platform and other software and hardware to ensure the compliance of data collection and the consistency of implementation;

2) The design of the data collection method should take into account the objective accuracy of the collected data, the safety of the collection method and the on-site feasibility at the same time, so as to avoid the use of improper collection methods to affect industrial production and operation;

3) In the process of data collection, authenticate and authenticate the data interface;

4) Monitor and audit data collection behavior, and promptly alert once abnormal behavior is found.

4.2.2. Data transmission encryption
Encrypted transmission methods include (1) One-way data transmission: Use unidirectional isolation means for data transmission, and the isolation strength should comply with the border security
strategy. (2) Encrypted transmission: Encryption technology and other methods should be used to ensure the integrity, confidentiality and availability of data during transmission. The selection of encryption algorithms and products should comply with relevant national regulations. At the same time, application scenarios, transmission methods, data scale, and efficiency requirements should be considered. (3) Security protocol or dedicated link transmission.

4.2.3. Data security monitoring
Security risk monitoring should establish an industrial Internet data security monitoring mechanism, including the establishment of industrial Internet data security monitoring technical requirements and implementation of security monitoring technical measures, etc.; real-time monitoring, analysis and early warning of industrial Internet data security risks through the construction of systems and platforms can be intuitive. Presents the industrial Internet data security situation, risk threats, flow paths and distribution of important sensitive data. In detail, the security monitoring platform can collect relevant data from the industrial Internet, industrial cloud and cross-border network entrances and exits, conduct preliminary analysis and screening of the data, and integrate industrial-related real-time monitoring data, asset management data, production operation data, and production management data. Operation and management data, equipment information, key parameters of industrial control equipment and other industrial sensitive data are screened out. And use blacklist, whitelist and greylist strategies to filter the data, and perform analysis and early warning. Common data security monitoring can be divided into sensitive information monitoring according to threat types. Malicious code propagation monitoring, vulnerability exploitation and attack monitoring, Data security monitoring strategy framework as shown in Fig 2.

(1) Malicious code propagation monitoring
By adding the characteristics and rules of malicious code such as Trojan horses and botnets of the Internet and industrial control networks into the data security monitoring system, it is possible to monitor the activities of malicious codes in various regions across the country, so as to understand and discover the master and controlled terminals in time, To provide effective technical support for dealing with these security threats.

(2) Vulnerability exploitation and attack monitoring
By adding the feature rules of computer systems and industrial control system vulnerabilities to monitor system, you can monitor the vulnerability attacks and find attacks on important systems as well grasp the situation of network attacks across the country. It can effectively identify typical industrial control vulnerabilities such as denial of service vulnerability (CVE-2013-2784) and buffer overflow vulnerability (CVE-2014-0768) of control devices such as PLCs, and generate warning messages.
4.2.4. **Data desensitization**

The sensitivity of the data should be evaluated, and the data should be desensitized according to the evaluation results and actual application requirements to ensure complete data desensitization and the availability of data after desensitization.

Appropriate desensitization methods should be selected according to data characteristics, security levels, application scenarios, etc., and at the same time, data desensitization effectiveness evaluation methods should be established to ensure the effectiveness of data desensitization.

An automated desensitization mechanism for sensitive data should be established, and desensitization methods such as anonymity, generalization, suppression, interference, randomness, and encryption should be used to desensitize data to improve the efficiency of desensitization.

4.2.5. **Data storage**

High-level encryption technologies such as complete encryption are used to realize the confidentiality, integrity and availability of stored data. At the same time, effective key management mechanisms and technical means are established to realize the full life cycle safety management of keys. Perform regular data backup and recovery tests to ensure timely, complete and accurate data recovery; according to the importance and sensitivity of the data, provide data local and remote disaster backup and recovery functions, formulate corresponding backup cycles, and store the backup media off-site; Provide hardware redundancy and thermal redundancy for main network equipment, communication lines and data processing systems to ensure high availability and good scalability of the system, and to ensure data recovery.

4.2.6. **Data archiving and destruction stage**

Data destruction and disposal should establish a data deletion mechanism and adopt technical means such as hard disk formatting and multiple erasing to ensure that the data is completely destroyed, leaving no traces, and cannot be restored. Effectively clear the cache data of the data import and export channel to ensure that the data involved in the import and export process will not be maliciously restored. The storage space with data is completely cleared before being released or reallocated to prevent illegal and malicious restoration.
References

[1] Industrial Internet Consortium: The Industrial Internet of Things Volume G4: Security Framework Version 1.0, 2016-September-26 http://www.iiconsortium.org/IISF.htm

[2] Zhang Xue ying, Chen Xue hong, Yang Shuai feng, “Research on the status quo of industrial internet data security standards in China,” J.Cyberspace Security vol.10, pp. 86–92, Oct 2019.

[3] Fan Peiru, Wang Chonghua, “Attack and protection of industrial internet platform from the perspective of data security”, J. Vol.11 No.2 Feb.2020, pp.75–80.

[4] Industrial Internet Industry Alliance, White Paper of Industrial Internet Platform [Z].2019.

[5] Industrial Internet Industry Alliance, Industrial Internet Security Framework [Z].2018.

[6] IoT Security Maturity Model: Description and Intended Use, IIC:PUB:IN15:V1.0:PB:20180409, 2018-April-09

[7] ISO 7498-2:1989.Security Architecture: General description of security services and related mechanisms covering secure communications between open systems.