DOI: 10.5281/zenodo.21988026

Physical Sensing Carrier Linked with Digital Storage to Complete Massive Information Collection Application

Haocheng Yuan

College of Physical Science and Engineering Technology, Guangxi University, Nanning 530004, China

Abstract: ***With the rapid development of the Internet of Things, industrial sensing, and smart city construction, the monitoring and collection data generated by various physical sensing devices are exploding. The massive, multi-source, and heterogeneous information poses great challenges to traditional data collection models. This article focuses on the linkage mechanism between physical sensing carriers and digital storage systems, elaborating on the basic theoretical system of sensor information collection, and analyzing the core problems in current software and hardware linkage, data processing, storage adaptation, and practical applications. Propose targeted optimization strategies and application solutions, build an efficient, stable, and secure massive information collection system, and provide technical support and practical reference for big data collection and resource utilization in industrial monitoring, environmental perception, intelligent operation and maintenance, and other fields.***

Keywords: Physical sensing carrier, Digital storage, Data collection, Massive information, Linkage mechanism, Intelligent perception.

1. INTRODUCTION

In the era of digital economy, global perception, real-time monitoring, and dynamic traceability are the main requirements for the intelligent upgrading of various industries. Various types of physical sensing carriers such as temperature, humidity, pressure, images, and displacement have become key terminals for the digital transformation of information in the physical world. The efficient collection, preservation, and integration of massive sensor information is the fundamental basis for conducting data analysis, intelligent analysis, and decision-making assistance. At present, with the large-scale deployment of sensing devices, data transmission has become fragmented, storage structures cannot meet the needs, and the coordination between software and hardware is poor, resulting in a large amount of data information collection work not being well implemented, data distortion, and waste of a lot of resources. Therefore, starting from the basic principles of the interconnection between sensing and storage, this article identifies the problems in industry applications and proposes practical optimization measures, with the aim of breaking down the barriers between physical sensing and digital storage, so that a large amount of sensing data can be standardized, efficiently, and scaled up for utilization.

2. CORE THEORY OF LINKAGE BETWEEN PHYSICAL SENSING AND DIGITAL STORAGE

The theory of integrating physical sensing carriers with digital storage information collection systems is mainly based on four aspects: sensor perception principles, data transmission protocols, distributed storage architecture, and data fusion processing. Physical sensing carriers mainly rely on sensing principles such as optoelectronics, mechanics, and electromagnetism to accurately convert physical signals into digital electrical signals, which is the foundation of information acquisition. The data transmission theory standardizes the encoding, encryption, and transmission time of sensor data to ensure that terminal data can be smoothly sent to storage terminals. Digital storage relies on the theories of distributed, cloud storage, and edge storage architecture to implement hierarchical storage and systematic management of massive data. The theory of data fusion can clean, normalize, and integrate heterogeneous sensor data from multiple sources. The four theories support each other and form a complete technical system for the digital collection of physical information, which is the main support for large-scale industry applications.

3. EXISTING PROBLEMS IN THE INTEGRATION OF PHYSICAL SENSING AND DIGITAL STORAGE FOR MASSIVE INFORMATION COLLECTION

3.1 The Heterogeneity of Data Collection in Sensor Terminals is Prominent, and the Standardization Level of Data Collection is Low

At present, there are various types of physical sensing carriers in the industry, and different manufacturers and models of temperature, vibration, image, and gas sensors have inconsistent acquisition frequencies, data accuracy, encoding formats, and output protocols. Some aging sensing devices use traditional proprietary protocols, while new smart sensors use standardized IoT protocols. When different types of terminals work simultaneously, the large amount of data structures generated are loose and each has its own characteristics. At the same time, most scenarios do not have a unified data collection standard, and situations such as data duplication, missing, and anomalies occur from time to time. When collecting data, a large amount of manual preprocessing is required, and automated data collection is difficult, greatly reducing the speed of collecting massive amounts of information. It cannot directly adapt to the storage standards of large-scale digital storage systems.

3.2 Insufficient Coordination Between Sensing and Storage Linkage, Resulting in Data Transmission Link Lag and Distortion

Both physical sensing terminals and digital storage systems suffer from the phenomenon of "separation of acquisition and storage", lacking a collaborative mechanism. On the one hand, most small and medium-sized application scenarios use traditional single point transmission methods, where sensing devices continuously transmit massive amounts of data to multiple locations. During peak periods, data congestion, delays, packet loss, and other issues may occur, and high-frequency real-time data collected cannot be uploaded to storage terminals in a timely manner. On the other hand, there is a mismatch between edge sensing terminals and cloud based digital storage architecture, with no dynamic scheduling mechanism formed. Low priority redundant data occupies most of the storage bandwidth, and the transmission of core monitoring data is also slow. In addition, the signals of some wireless sensing devices are unstable, and the data is prone to interference and distortion during transmission, resulting in incomplete and inaccurate collected data.

3.3 Insufficient Adaptability of Digital Storage Architecture and Low Efficiency in Managing Massive Data

Due to the PB level massive dynamic data generated by sensing devices, traditional centralized storage methods have obvious shortcomings, such as difficulty in expanding storage capacity and limited read and write speeds, which cannot meet the needs of high-frequency and large-scale storage of sensing data. Most storage systems do not optimize the temporal and periodic characteristics of sensor data. All data is stored in one place, regardless of type or level, resulting in invalid redundant data occupying a large amount of storage space, and accessing and retrieving core data is extremely difficult. At the same time, the security mechanism for storing sensor data is not sound, and there are no hierarchical encryption, permission control, and data backup measures in place for storing massive amounts of multi-source data. There is a risk of data leakage, loss, or tampering. Moreover, there are data interface issues between the storage system and subsequent analysis and application platforms, which prevent the collected large amount of data from being converted into usable data resources in a timely manner.

4. THE IMPLEMENTATION STRATEGY OF LINKING 3 PHYSICAL SENSING CARRIERS WITH DIGITAL STORAGE TO ACHIEVE MASSIVE INFORMATION COLLECTION

4.1 Unified Data Collection Standards to Achieve Standardized Collection of Sensor Information

In terms of the problem of heterogeneous and low standardization of sensor data, it is necessary to establish a unified collection and preprocessing standard, and create a standardized data aggregation front-end system. Firstly, it is necessary to standardize the transformation of terminal devices, eliminate inefficient sensors with outdated protocols, low accuracy, and high energy consumption, and upgrade existing in use devices through protocol adaptation. Unified IoT transmission protocols such as MQTT and Modbus should be used to unify data encoding formats, collection frequencies, and output dimensions, fundamentally solving the problem of chaotic data structures. In addition, the front-end data pre-processing module is set up, and the edge computing gateway is deployed to conduct real-time cleaning, de duplication, supplementation, standardization and other processing of various sensor data, eliminate abnormal and redundant data, and unify the data formats output by different sensors. Finally, establish an industry adapted data collection standard system, which specifies the collection indicators, transmission timing, and storage standards for sensor data in various scenarios, so that all sensor data can be automatically stored according to a unified standard, and a large amount of information can be automatically archived without manual intervention. On this basis, a device data ledger and dynamic matching mechanism can be established to perform standardized testing on new sensors before integration, ensuring that the parameters of all network connected terminals meet the requirements and the protocol is unified. In addition, data quality verification rules are added to set threshold verification and temporal continuity verification, ensuring the quality of data collection at the source and guaranteeing the normalization and continuous storage of massive data.

4.2 Optimize the Linkage Transmission Mechanism and Break Down Barriers to Data Transmission Across the Entire Link

Establish a dual pronged linkage system of "edge transmission and cloud scheduling" to eliminate the lag caused by low collaboration between sensors and storage and data transmission distortion. Partition and stratify the transmission link, using wired primary transmission and wireless backup transmission modes to ensure stable data transmission in complex industrial and outdoor environments, and using low-power wireless transmission modules to reduce bandwidth consumption in civilian lightweight scenarios. Build a dynamic bandwidth scheduling system to dynamically allocate transmission resources based on data priority, setting important and urgent data as high priority to ensure timely transmission. For general monitoring data, staggered transmission is used to avoid network congestion. In addition, a data transmission verification and retransmission mechanism will be added. After the terminal uploads data, the storage system will automatically generate data integrity verification, generate automatic retransmission requests for missing or distorted data, and create a transmission log traceability system to ensure the accurate and complete transmission of a large amount of sensor data, achieving instant connection between sensor acquisition and digital storage. Real time monitoring of the link status is implemented simultaneously, with routine monitoring of important parameters such as network latency, packet loss rate, and signal strength. Once there is a fluctuation in the link, an alarm is immediately issued and a link switching operation is initiated. For large-scale sensor networking, adopting lightweight transmission methods such as sharded transmission and bulk packaging can reduce the burden of high-frequency data transmission, thereby enhancing the stability and speed of massive data transmission.

4.3 Iterative Upgrade of Storage Architecture to Meet the Efficient Management Needs of Massive Data

Based on the temporal, massive, and real-time characteristics of sensor data, upgrade and optimize the digital storage architecture, and create a hierarchical and classified intelligent storage system. Abandoning the traditional centralized storage model, we have created a fusion architecture of "edge storage+distributed cloud storage". In the short term, real-time sensor data is stored in edge terminals, which can be quickly accessed. In the long term, massive historical and important data is transmitted to the distributed cloud storage cluster, which can be infinitely expanded. Storage space and computing resources can be expanded at any time according to data increments. Create a hierarchical storage system for data, dividing it into real-time data, regular data, and archived data based on their value and timeliness. Adopt different configuration strategies for high-speed cache, regular storage, and cold storage to significantly reduce storage energy consumption and expenses, and enhance data retrieval speed. Improve the storage security system, implement layered encryption and storage of core sensor data, set up multi-level access permissions, establish an automatic backup and disaster recovery system, carry out data incremental backup work on time, prevent data loss or leakage, and ensure the security and stability of massive aggregated data. To adapt to the characteristics of massive data updates, intelligent data lifecycle management can be deployed, which will automatically perform data migration, compression, or destruction operations after changes in factors such as time and access times. In addition, it is necessary to optimize the temporal database and create a dedicated indexing mechanism for sensor data, which significantly accelerates the speed of querying, statistics, and analysis of massive temporal data, thus solving the problems of slow and low adaptability in traditional database queries.

4.4 Building an Integrated Linkage Platform to Achieve Data Collection and Resource Utilization Applications

In order to completely solve the problem of disconnection and low utilization between sensors and storage, an intelligent linkage management platform that integrates data collection, transmission, collection, storage, analysis, and application has been created to achieve full process implementation. The platform consists of five main parts: terminal management, data collection, storage management, data application, and security monitoring. The terminal management module can timely grasp the operating status and collection parameters of all physical sensing devices, and can carry out centralized control of devices and issue alarm messages in case of faults; The data collection module summarizes and organizes the preprocessed data from the front-end, and then automatically implements the process of batch storage and classification saving, generating standardized data ledgers; The storage management module monitors the operation status of the storage cluster, automatically achieving tasks such as capacity expansion, data migration, and garbage data removal to ensure the normal operation of the storage system. The platform's open and standardized data interface can be integrated with downstream systems such as data analysis, intelligent decision-making, and visual display. After obtaining a large amount of sensor data, deep level mining work can be carried out to achieve application purposes such as equipment status monitoring, environmental development trend prediction, and fault prediction. In addition, establish a full process monitoring and early warning mechanism to monitor the entire process of data collection, transmission, storage, and application in real time. Once abnormalities are detected, alarm signals will be issued, and the linkage collection system will be continuously improved to transform massive sensor information from collection and collection to resource utilization. To enhance the practicality of the platform, mechanisms for data operation and collaborative management can be established. Different data collection templates and permission groups can be set up in various industries to meet the needs of various application scenarios such as industry, environmental protection, and security. Embedding an AI intelligent analysis engine into the existing system to perform clustering analysis, trend prediction, and anomaly tracking on the collected large amount of time-series sensing data, transforming passive data storage and aggregation into active data-driven, and adding a data traceability audit mechanism to complete the logging process from data collection to application completion. This effectively strengthens the practical application level of massive information and significantly improves the compliance standards and regulatory management of the industry.

5. CONCLUSION

The efficient linkage between physical sensing carriers and digital storage is the key to the digital collection and intelligent application of massive physical information, and it is also an important support for various industries to achieve intelligent transformation. This article summarizes the main theories of sensing and storage linkage, accurately identifies pain points in industries such as low data standardization, poor linkage, and storage mismatch, and provides comprehensive and feasible improvement strategies from four aspects: standard collection, chain improvement, architecture update, and platform creation. By establishing a complete sensor network technology and management system, the collection speed, quality, and security of massive sensor information can be greatly improved, fully realizing the value of sensor big data and providing support for digital transformation in various fields such as industrial intelligence, smart city construction, and environmental monitoring.

REFERENCES

[1] Liu Yongfeng, Zhang Chao Physical simulation and impact analysis of large-scale tailings dam failure based on multi-source sensor monitoring [J]. Journal of Henan University of Technology (Natural Science Edition), 2025, 44 (6): 64-74

[2] Yang Yingqi, Li Keyi, Lin Guiyang, etc Medical grade development of flexible respiratory sensing technology: Medical physics implementation pathway and clinical lung function data validation [J]. Acta Physica Sinica, 2025, 74 (14): 323-339

[3] Guo Ruimiao Research on the Application of Digital Information Storage in Digital Libraries [J]. Information Recording Materials, 2024, 25 (12): 154-156

[4] Xu Huaisheng, Xu Miaomiao Research progress on digital information storage technology for nucleic acid molecules [J]. Computer Knowledge and Technology, 2024, 20 (32): 77-79

[5] Zhou Jingyi Research on the Theoretical Data Carrying Capacity of Digital Information Storage in Microfilm [J]. China Archives, 2024, (7):20-21.

[6] Liu Wenlin Design of Multi Cloud Storage Platform for Massive Image Information Supported by Digital Technology [J]. Automation and Instrumentation, 2024, 39 (2): 135-139