J. Highw. Transp. Res. Dev. 2025, 19(1): 33−38

Journal of Highway and Transportation Research and Development

Home page: www.sciopen.com/journal/2095-6215

Data collection and transmission strategies for structural safety monitoring of large-scale infrastructure

Feng Zhou**

Environmental Protection Company of CCCC Tianjin Dredging Co., Ltd., Tianjin 300461, China

Abstract: With the goal of safety monitoring of large-scale infrastructure structures,Systematically sort out the indicators that affect structural safety, establish a system for monitoring structural safety based on key indicators, an overall plan for massive real-time data acquisition and transmission of large-scale infrastructure structure safety monitoring systems was designed, and the data flow of each monitoring item of the system was sorted out, proposed a real-time data collection strategy and overall architecture.By mining, cleaning, and preprocessing heterogeneous data from multiple sources, high-quality data storage can be achieved, which not only effectively realized the shared use and centralized management of data collected by infrastructure monitoring systems in various places, but also provided experience for the later data analysis of large-scale infrastructure structure monitoring systems support.

Keywords: Bridge engineering; Infrastructure; Structural monitoring; Multi-task coupling; data; Real-time data flow

1 Introduction

As infrastructure construction experiences vigorous growth, the safe operation of these vital structures has become a top priority. In response, domestic structural safety monitoring systems for large and medium-sized infrastructure projects have undergone rapid development. Latest statistics reveal that there are now hundreds of such monitoring systems deployed across the country, amassing nearly one million valid data points and pieces of information every day. These systems are instrumental in the ongoing maintenance management, real-time data analysis, risk assessment, and early warning for infrastructure. To align with the requirements of modern infrastructure management, various cities in China have established their own infrastructure management systems tailored to their specific needs. Shanghai has pioneering a large-scale infrastructure database utilizing the power of the Internet, complete with diverse information query and statistical tools, as well as data management models designed to meet user specifications [1]. Anhui Province has also harnessed information mining technology to develop a comprehensive data collection and transmission system for signal acquisition, processing, caching, and transmission [2]. Liu et al. [3] developed an acquisition conditioner that leverages the 1 588 clock synchronization protocol to enable precise synchronization for the acquisition of high-frequency signals. Peng Hong and colleagues [4] utilized National Instruments (NI) technology combined with LABVIEW to perform synchronous data acquisition and transmission, and they also implemented data screening and cleaning procedures.Despite these technological advancements, the data acquisition and transmission methods used in current infrastructure structural safety monitoring systems are inconsistent. There is no unified, cross-regional information data interface support, which impedes interoperability among different systems. A significant gap exists in the real-time performance, effectiveness, and efficiency of data collection, with the data being scattered across various locations. This fragmentation prevents the sharing and effective use of infrastructure structural data, presenting substantial challenges for subsequent data management and the successful implementation of information mining techniques [5].

To facilitate more convenient and efficient management and maintenance of infrastructure by relevant departments, and to maximize the utility of the diverse data collected on-site, it is essential to conduct research into strategies for data collection within the context of real-time systems designed for infrastructure structural safety monitoring.

2 Large infrastructure structure safety monitoring system composition

The current implementation plan for establishing a structural health monitoring system for large-scale infrastructure is shown in Fig. 1, which mainly includes an automated sensing subsystem, a data acquisition and transmission subsystem, a data preprocessing subsystem, a database subsystem, and an evaluation subsystem.

(1) Automated sensing subsystem. The automation sensor subsystem is a hardware system that continuously monitors structural response and environmental information, and transmits them to the data acquisition and control subsystem in the form of physical quantities such as light and electricity.

(2) Data collection and transmission subsystem. This subsystem consists of hardware and software systems, which serving as a

Received: 26 July 2024; **Revised:** 21 September 2024; **Accepted:** 28 December 2024
** *Corresponding author*: 94175213@qq.com
DOI: 10.26599/HTRD.2025.9480052

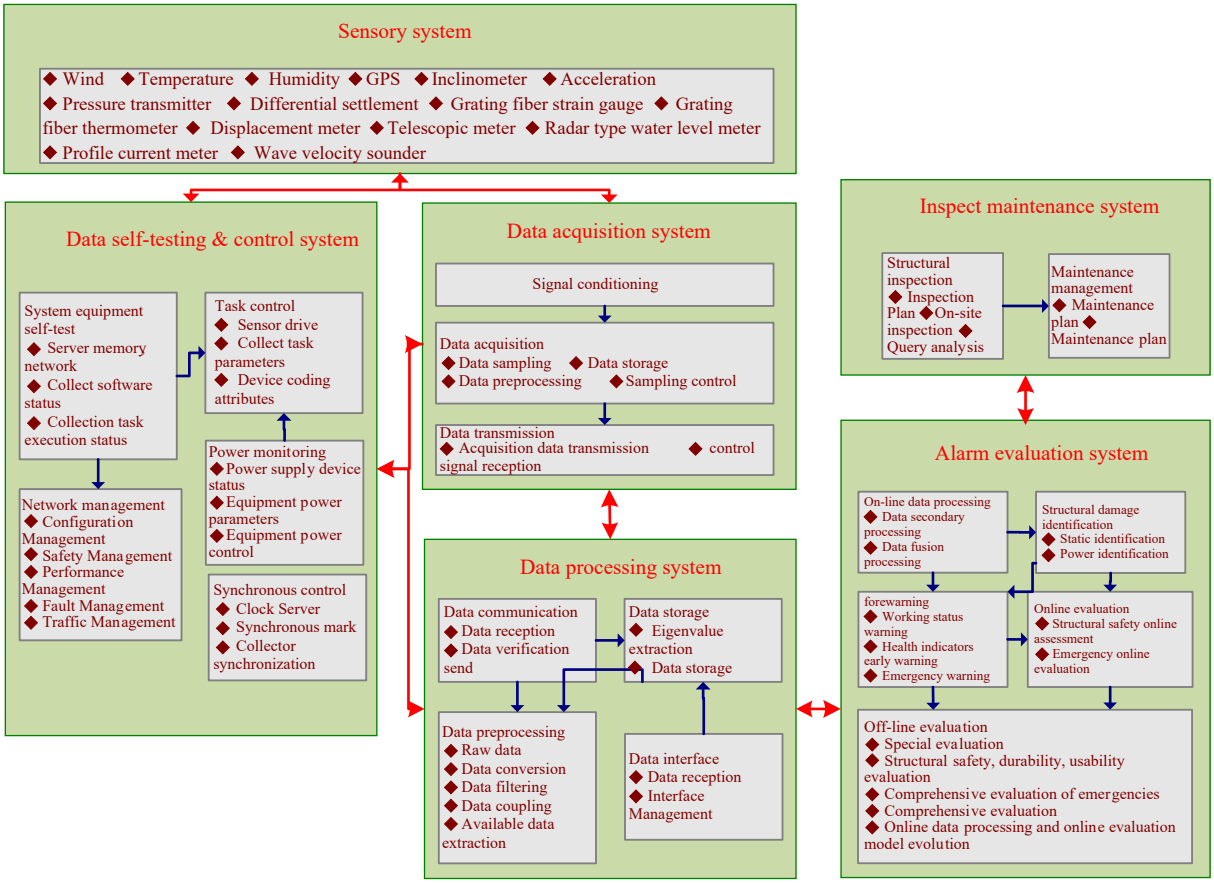

Fig. 1 Architecture diagram of infrastructure structural health monitoring system.

bridge between the sensor subsystem and the data management and processing subsystem. Through this subsystem, the digital signals such as light and electricity transmitted by the sensor subsystem can be converted into analog signals, realizing the transmission channel between the sensor subsystem, the acquisition and control subsystem, and the data management and processing subsystem, and stored in the data management and processing subsystem.

(3) Data management and processing subsystem. This system is responsible for the data management and processing functions of the structural health monitoring system. The collected data is preprocessed by hardware on analog signals, and then preprocessed by data processing software (including filtering, extraction, mining, data coupling between channels, and conversion) to complete data filtering and mining.

(4) Database subsystem. The core of this system is the database, mainly used for storage and management of historical data. According to the functions and requirements, a database cluster is adopted to store data and solve the problem of difficult data reading.

(5) Structural evaluation subsystem. This subsystem is mainly composed of multiple software modules such as damage identification software, model correction software, structural safety assessment software, reliability calculation software, and real-time warning software. It is the core part of the functional implementation of the structural health monitoring system.

3 Strategy and overall architecture of data collection

The real-time collection and transmission of on-site data is the hub and core of the entire infrastructure safety monitoring system. The main task is to complete the acquisition, conversion and transmission of raw data based on physical sensors [6], which is the link between on-site collection and it is a key link in the equipment and backend data processing center, so the system is required to have high stability and self-repairability. We use today's mainstream Linux platform as the operating platform for the collection software.

3.1 Overall strategy for data collection

The Linux operating system is a multi-tasking and multi-user operating system with good real-time response performance and multi-tasking capabilities, as well as high reliability and scalability [7]. In view of the characteristics of this system, the data collection software is implemented based on the Linux platform, which can achieve good real-time and high accuracy of data monitoring. At the same time, the system has good reliability, stability and anti-interference characteristics. A data collection and transmission server is arranged at each outstation and monitoring center [8]. The two servers serve as backup for each other, greatly improving the reliability of the system (Fig. 2).

3.2 Overall data collection functional analysis

The collection terminal software works in the Linux environment and mainly completes the collection and storage of signal data and physical data with the support of terminal hardware collectors and sensors. Users generally do not interact with it directly, but it provides a series of standard interfaces and commands to interact with the control terminal, monitoring terminal and data storage terminal where the user is located. The functions planned to be provided by the collection terminal software are as follows [9]:

① Standardize physical data collection methods and byte protocols, and unify collection strategies, real-time specifications and interfaces.

② Complete data collection and summoning mechanism, working condition query and statistical functions of collection points.

③ Configuration information loading function reads the system information conFig. d in the configuration file, and dynamically

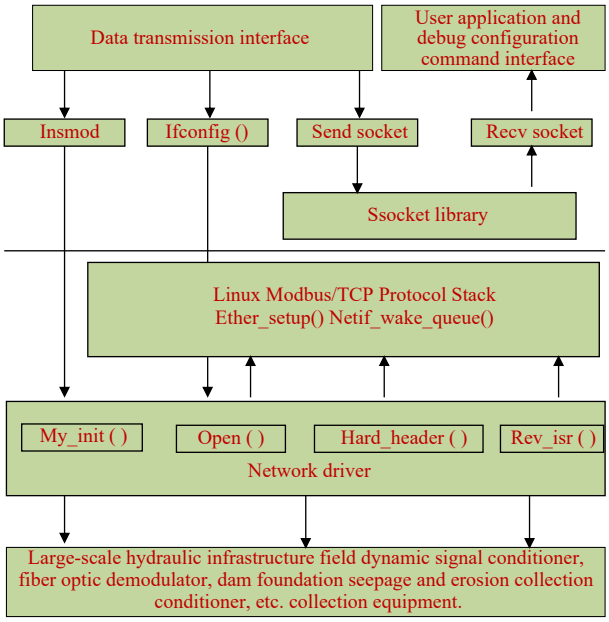

Fig. 2 Overall architecture diagram of data collection.

generates threads for data collection and processing, data storage areas, etc. based on the configuration information.

④ Precise timing function, data summoning, thread scheduling, data processing and other tasks provide precise time information.

⑤ The data summoning function relies on the precise timing function to periodically summon data from the collection point based on the read configuration information.

⑥ The data receiving and processing function realizes receiving and processing of data sent from the collection point, and puts the processing results into the corresponding data storage area.

⑦ The temporary data storage function stores the collected data and records the corresponding information for use by the transmission module.

⑧ Query and record the communication conditions of the collection point equipment, and publish the communication conditions to the network in real time for use by interfaces, background and other systems.

At the same time, the configuration file of the collection software platform supports the setting of collection tasks for the collection terminal. Mainly include:

① Acquisition module sampling rate.

② Attributes of each collection channel (corresponding to the ID of the collection channel, which is the unique identifier of the collection signal in the entire monitoring system).

③ Collect signal thresholds (maximum and minimum values) and fault flags.

④ Unit conversion coefficient (the conversion relationship between the collected electrical signal data and the original physical signal, based on this relationship, the electrical signal data is converted into the original physical signal data).

⑤ Data storage format and storage mode.

The task settings of the collection module do not support non-static real-time changes. Before changing the settings of the collection module, the current collection task must be stopped, and then the commands related to task and mode changes must be uploaded and updated to subsequent new collection tasks [10]; The collection terminal sends the continuously collected data in units of data packets (a data packet contains signal data continuously collected over a period of time and some corresponding signal information) to the monitoring terminal through the UDP port; the collection terminal sends the data storage task The specified data packet is stored locally in the form of a text file; the collection terminal sends the status conversion of each internal module, engine and interface, error information and other contents to the control and monitoring terminal through broadcast, and organizes the error information into an error log Stored locally; each acquisition terminal uses a 1 588 clock server to accurately time the vibration signal of the bridge. Synchronous acquisition meets strict phase synchronization requirements between vibration signals at each point.

4 Technical analysis and implementation of data collection

4.1 Data collection technical considerations

The acquisition system is installed and run on the infrastructure site, communicates with on-site collectors and other equipment, collects various parameter data, and transfers the data to the real-time database system installed in the control room [11]. Its technical characteristics are as follows:

① Using concurrent programming technology, the system can communicate with multiple collectors in parallel at the same time. Ensure high-speed and efficient data collection to meet the requirements of the monitoring and analysis system for data collection frequency.

② Adopt an abstract acquisition-driven model. This makes the program architecture highly adaptable to different types of collection devices. When adding a new collector, just add a driver module instance.

③ The collection system provides external data in the form of services and can support simultaneous access by multiple clients. In this way, the real-time library system can read data from the acquisition system concurrently with multiple tasks, improving the data transmission speed. On the other hand, other programs, such as test tools, can also read data simultaneously, making on-site debugging convenient.

④ The collection system provides runtime information to the outside world in the form of services. In this way, even though the collector is placed at the site of the infrastructure, its operation can still be monitored remotely.

⑤ The acquisition system has the function of local data storage, so that when the communication is abnormal, the data can be maintained for several days without loss, improving the stability of the system.

⑥ The verification technology to prevent packet loss is considered in the design of data transmission protocol of the acquisition system. Even in the case of wireless communication, it can still ensure the correctness of data transmission.

4.2 Overall data flow diagram

The monitoring data flow of infrastructure structure is very large, which not only requires very high real-time, but also has accurate time synchronization for high speed data such as acceleration. According to the acquisition method, the data acquisition system can be divided into modular-based acquisition module, grating fiber strain acquisition system, basic scour acquisition system, dynamic traffic load monitoring system and GPS spatial displacement acquisition system. The specific data flow is shown in Fig. 3 [12,13].

4.3 Data acquisition software structure design

In order to facilitate and standardize the debugging and use of the acquisition terminal software, each acquisition terminal uses a unified software architecture. Since the hardware configuration of each terminal is different and the physical signals collected are also different, when designing and implementing the data collection terminal software, the following aspects should be fully considered [14]:

① Standardizing shareware architecture.Each sub-station adopts the same standard and shareable software model, which facilitates the unified release of collection tasks and consistent reporting of

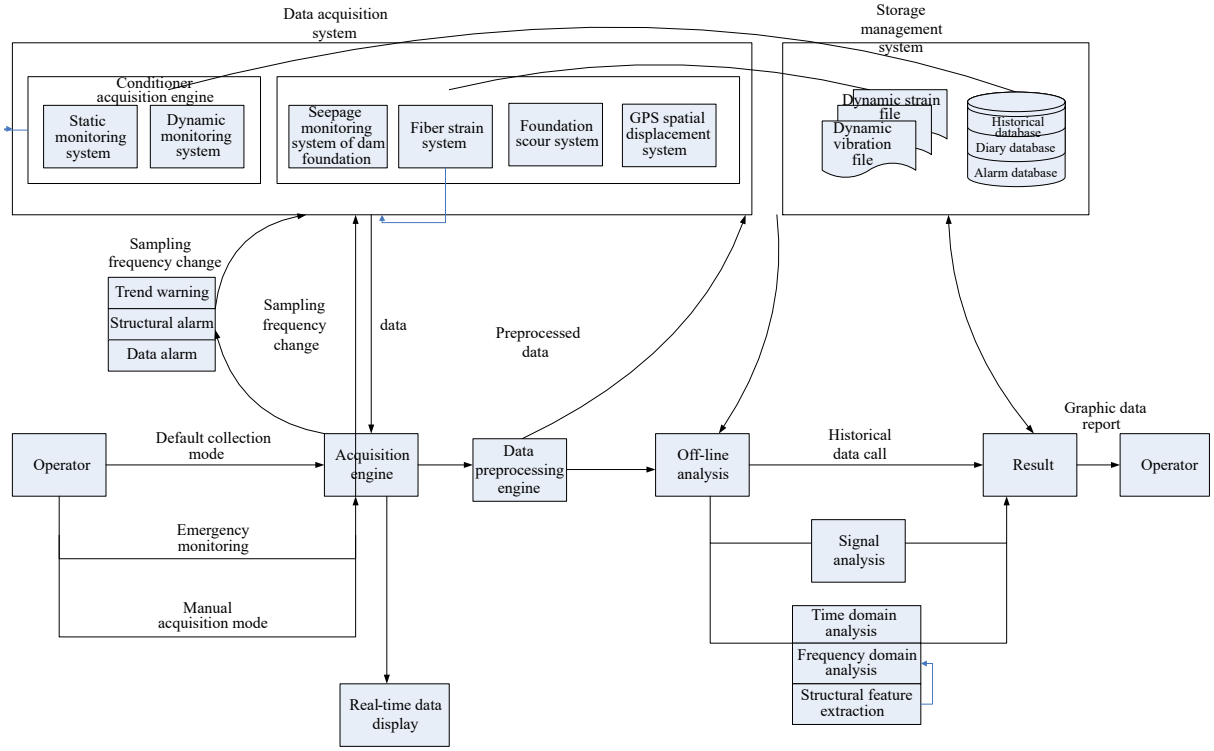

Fig. 3 Data flow diagram of infrastructure structure health monitoring system.

status. At the same time, it can more conveniently collect equipment status, obtain collection information, and unify the interaction protocol [15].

② Modular, scalable data collection capabilities.Since most of the hardware configurations of each collection terminal are different, and considering the possibility of adding or changing measuring points and adjusting the system in the future, the software of the data collection terminal must be highly modular to facilitate developers to add new measuring points, hardware [16].

③ Easy-to-use operation interface for maintenance and configuration.Since each data collection terminal is located relatively far away, it must have a convenient interface to enable developers and users to change system settings in a timely manner. In this way, in the early stages of system installation and debugging, when parameter settings need to be frequently changed, a large amount of debugging time can be reduced.

Based on the above considerations, the software structure block diagram shown in Fig. 4 is given below.

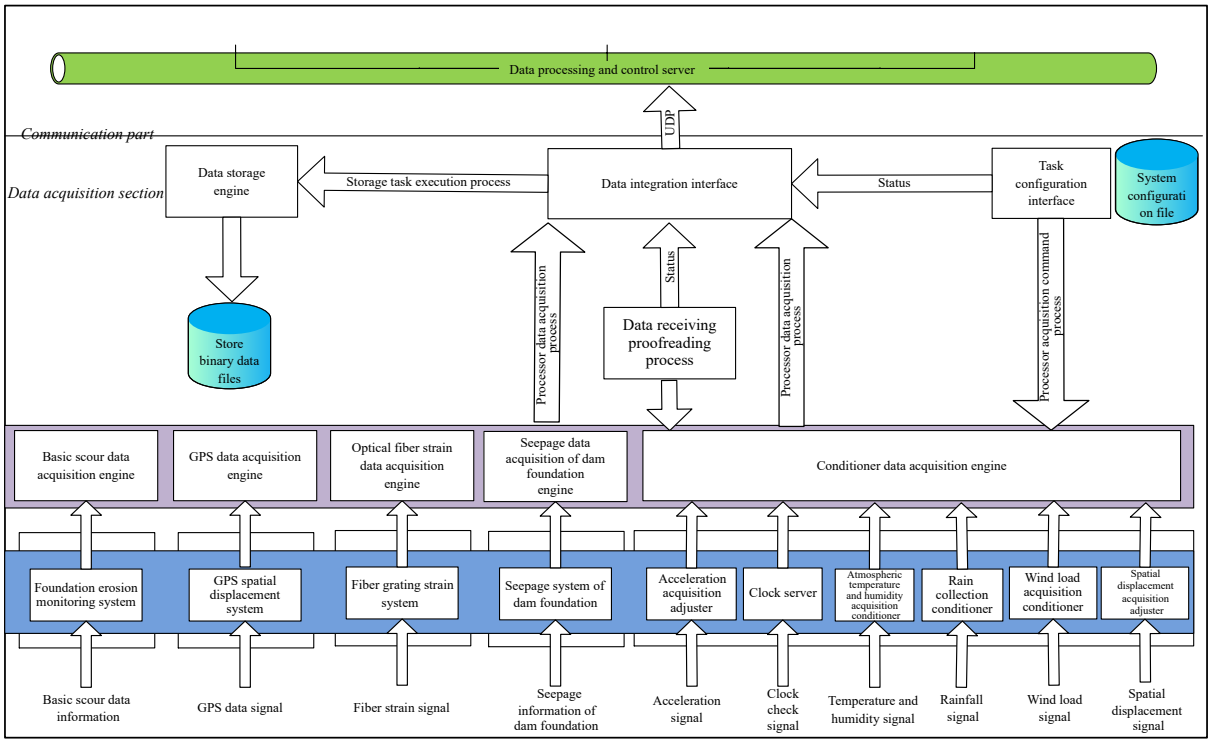

Fig. 4 Overall architecture diagram of data collection for infrastructure structure monitoring system.

4.4 Data processing methods

① Outlier removal [17]. For data within a specified time period, data points with absolute values greater than the threshold value are set as the mean or constant of their left and right data points. Firstly, the threshold value is determined by the magnitude of the absolute value, positive value, or negative value, and then it is judged whether it is a outlier. When a outlier is encountered, it is averaged.

② Filtering [18]. During the transmission of the measured signal, noise and interference signals are often mixed in due to external influences, and sometimes external interference can even cause signal distortion. Digital filtering can be implemented through certain algorithms to change the relative proportion of frequency components in the measured signal or remove certain frequency components. Therefore, a mobile smoothing filtering algorithm was added to analyze the characteristics of each component in the acceleration measurement values.

Select a continuous time period T and take the mathematical average of the sampled signal $a(t)$ during T period, where the mean is an approximate value of the gravity component. The noise $\delta(t)$ can be filtered out by low-pass filtering, and the approximate acceleration value can be obtained by subtracting it from the measured value. The key to the algorithm is the selection of time period T, which requires.

$$\frac{1}{f_{g\max}} \gg T > \frac{1}{f_{v\max}} * (2/3)$$

where, $f_{g\max}$ is the highest frequency of change of $ag(t)$, indicating $f_{g\max} = 0.007$ Hz; $f_{v\min}$ is the lowest value of the main response frequency of $ag(t)$. By experimental observation, taking $f_{v\min}$= 10 Hz, it can be determined that $T = 0.4$ s.

Set a array of length N during algorithm implementation, and resample data from the sampled samples at time intervals t, satisfying $t \times N = T$. The selection of N is related to the memory size of the processor. Every time a new data is extracted, it replaces the first extracted data in the original sequence and continues the calculation and processing.

③ Sliding average processing [19]. Temperature sampling T_i is performed every ten seconds. Assuming that six sampling values T_1, T_2, T_3, T_4, T_5, T_6 are calculated every 10 minutes, $T_{\max}$ and $T_{\min}$ are removed, and the arithmetic mean of the remaining four sample values is calculated as the observation value for that minute, which is defined as the instantaneous value $T_a = \frac{T_1 + T_2 + T_3 + T_4}{4}$. The highest and lowest temperatures for one hour and their corresponding occurrence times are selected from the temperature observations for each minute, and the temperature observation values are stored.

5 System application projects and fields

The data collection software adopts the mainstream Linux platform today. The Linux operating system is a multitasking multiuser operating system with good real-time response performance and multitasking processing capability, high reliability and scalability. The data preprocessing software has a large amount of data. Considering work efficiency, it adopts a C/C++ working environment and does not use any third-party libraries other than the standard C library, strictly controlling the system's memory usage.This solution has been implemented by software and applied to structural monitoring,operation and management of a large infrastructure. It can not only make up for the shortcomings of data collection, transmission, processing and management of the new system, but also significantly improve The completeness, accuracy and real-time nature of data collection can also be perfectly integrated into large-scale structural monitoring systems that have been used. The specific data flow is shown in Fig. 5.

Fig. 5 Field data acquisition and network monitoring center

6 Conclusion

The proposed system effectively addresses the challenges associated with the large volume of data collection, suboptimal data quality, and complex data transmission that currently plague the monitoring process of large-scale infrastructure projects. By harnessing information technology, it provides essential data for subsequent infrastructure management and maintenance, thereby facilitating targeted health assessments and enabling maintenance management personnel to perform maintenance activities with precision. However, the system's data processing capabilities are currently constrained, which typically results in the collection of raw data through conventional methods. Should further processing and sophisticated analysis be necessary, the use of additional software becomes imperative. The problem of rapidly recovering data with limited processing capabilities remains unresolved, posing a significant challenge. Moreover, traditional data collection methods store data locally on individual computers or servers, rendering it

vulnerable to security breaches, such as hacker attacks and viral infections, and compromising the integrity of data security.

In subsequent sections of this paper, we will introduce an enhanced data collection and computation platform that promises higher data quality. We will also outline rigorous data security management strategies and discuss approaches to further improve the management and exploitation of data within infrastructure monitoring systems.

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

Data Availability Statement

Some or all Data, models, or code that support the findings of available from the corresponding author upon reasonable request.

References

[1] Zan X,Ou GL. Transport infrastructure, internet and urban economic growth potential. Journal of Highway and Transportation Research and Development, 2023, 40(1): 261–270. DOI: 10.3969/j.issn.1002-0268.2023.01.030.

[2] Salvadori F, De Campos M, De Figueiredo R, et al. Monitoring and diagnosis in industrial systems using wireless sensor networks. 2007 IEEE International Symposium on Intelligent Signal Processing. Alcala de Henares: IEEE, 2007: 1−6.

[3] Ahuja K, Khosla A. A novel framework for data acquisition and ubiquitous communication provisioning in smart cities. Future Generation Computer Systems, 2019, 101: 785–803.

[4] Liu FL, Li N, Feng LP, et al Research and implementation of clock synchronization server based on IEEE1588. Shandong Automation Society, Shandong Institute of Automation, Shandong University of Technology, 2010.

[5] Sun LM , Zhang QW. Large-span bridges and their health monitoring systems in China. International Symposium on Integrated Life-cycle Design & Management of Infrastructure.2007.

[6] O'Connor SM, Lynch JP, Gilbert AC. Implementation of a compressive sampling scheme for wireless sensors to achieve energy efficiency in a structural health monitoring system. Nondestructive Characterization for Composite Materials, Aerospace Engineering, Civil Infrastructure, and Homeland Security 2013. San Diego. SPIE, 2013: 86941L. https://doi.org/10.1117/12.2010128.

[7] Gillen A , Waldman B. Linux in the Mainstream : Growing Deployment of Business-Critical Workloads. International Data Corporation (IDC), 2011, 11: 228468.

[8] Lu C. The design of anti-interference and reliability in MCU system. Instrumentation Technology. Instrumentation Technology, 2010. https://api.semanticscholar.org/CorpusID:113280617.

[9] Liu DM, Wang WJ, Yang ZF, et al. Study on Application Mechanism of Internet+Epoch Crowdsourcing Traffic Big Data. Journal of Highway and Transportation Research and Development, 2018, 35 (7): 120−127. DOI: 10.3969/j.issn.1002-0268.2018.07.017.

[10] Patton AD .A Probability method for bulk power system security assessment - iii - models for stand-by generators and field data collection and analysis. IEEE Transactions on Power Apparatus & Systems, 2007, 91(6): 2486−2493. DOI:10.1109/TPAS.1972.293428.

[11] Marshall S. Optimarin applies big data to ballast water management. Naval Architect, 2021 (July/August).

[12] Li YY, Li M, Tian T, et al. A strain sensor based on core-distortion long-period fiber grating. IEEE Photonics Technology Letters, 2023, 35(24): 1339–1342.

[13] Zhou YB, LI R, LIU Haizheng, et al. Simulation of Random Traffic Flow and Analysis on Bridge Fatigue Response Based on CA. Journal of Highway and Transportation Research and Development, 2020, 37(5): 83–91,122.

[14] Zhao YZ, Hu CD, Sheng P, et al. Design of timing system software on eAST–NBI. Journal of Fusion Energy, 2013, 32(5): 557–560.

[15] Saha B, Adl-Tabatabai AR, Jacobson Q. Hardware acceleration for a software transactional memory system: US20060349787. US2007186056A1. [2025-01-23]. doi:US8521965 B2.

[16] Shen YH. Information Monitoring of animal husbandry industry based on the internet of things and wireless communication system. Computational and Mathematical Methods in Medicine, 2022: 8794044. https://doi.org/10.1155/2022/8794044.

[17] Parra-plazas J, Gaona-garci AP, Plazas-nossa L. Time series outlier removal and imputing methods based on colombian weather stations data . Environmental Science and Pollution Research, 2023, 30(28): 72319−72335.

[18] Chang ZW, Gao QH, Monti G, et al. Selection of pulse-like ground motions with strong velocity-pulses using moving-average filtering. Soil Dynamics and Earthquake Engineering, 2023, 164: 107574.

[19] Guo YJ, Fang Z, Li Q, et al. Design of water surface ultrasonic array imaging experimental system. Experimental Technology and Management, 2023, 40(12): 58–65, 91.