Computing&AIConnect
2025, Vol. 2, Cite as: doi:10.x/journal.x.x.x

SCIFINITI PUBLISHING

OPEN ACCESS

Review Article

Secure and Privacy-Preserving Data Management in Train Coupling/Decoupling Scenarios: A Comprehensive Review and Future Perspectives

Zhihang Zhang, Feng Wang, and Peng Li

School of Automation and Intelligence, Beijing Jiaotong University, 100044, China,

Abstract

This paper systematically investigates the data privacy protection strategies in railway transportation systems, with a particular focus on the unique requirements of train coupling/decoupling scenarios. In such dynamic and semi-open environments, existing passive security mechanisms and centralized architectures often fail to ensure secure data exchange and privacy preservation, particularly under limited computational resources. To address these challenges, we review decentralized data-sharing technologies and privacy-preserving computation methods suitable for heterogeneous train networks. Furthermore, we propose a blockchain-based asymmetric encrypted storage framework and a collaborative computing architecture based on federated learning, both tailored to the operational constraints of modern high-speed trains. Our approach integrates container virtualization, secure consensus protocols, and differential privacy techniques to enable traceable, tamper-proof, and privacy-aware data processing. Finally, this paper outlines future research directions concerning quantum-resistant security architectures and adaptive privacy mechanisms that can support the evolving needs of intelligent railway systems.

Keywords

intelligent railway transportation system; data privacy protection; train coupling/uncoupling; blockchain

I Introduction

Rail transportation has been widely adopted owing to its high transport capacity, safety, cost-effectiveness, and environmental sustainability. It is foreseeable that railway transportation is expected to remain vital to the global transportation systems, addressing future challenges such as longer transport distances, increasing passenger flows, and the growing demand for sustainable mobility. China holds a leading position in the railway area, particularly in high-speed railways. Statistics indicate that at the end of 2023, China's total railway operational mileage has reached 159,000 kilometers, with the HSR covering 45,000 kilometers. Additionally, over 4,300 electric multiple units (EMU) have been in service, accounting for 68.8% of the world's total HSR mileage. According to the Chinese government plan, the total railway operational mileage is expected to expand to 165,000 kilometers by 2025, with 50,000 kilometers of HSR.

The rapid development of high-speed rail raises higher standards for the safety, reliability, and operational efficiency of EMUs. As the central system of EMUs, the

existing train network control system is a typical distributed onboard control system that facilitates seamless integration between onboard equipment and ground infrastructure. It enables vehicle data communication, train-to-ground information exchange, and train control functionalities. The system architecture is illustrated in Fig. 1.

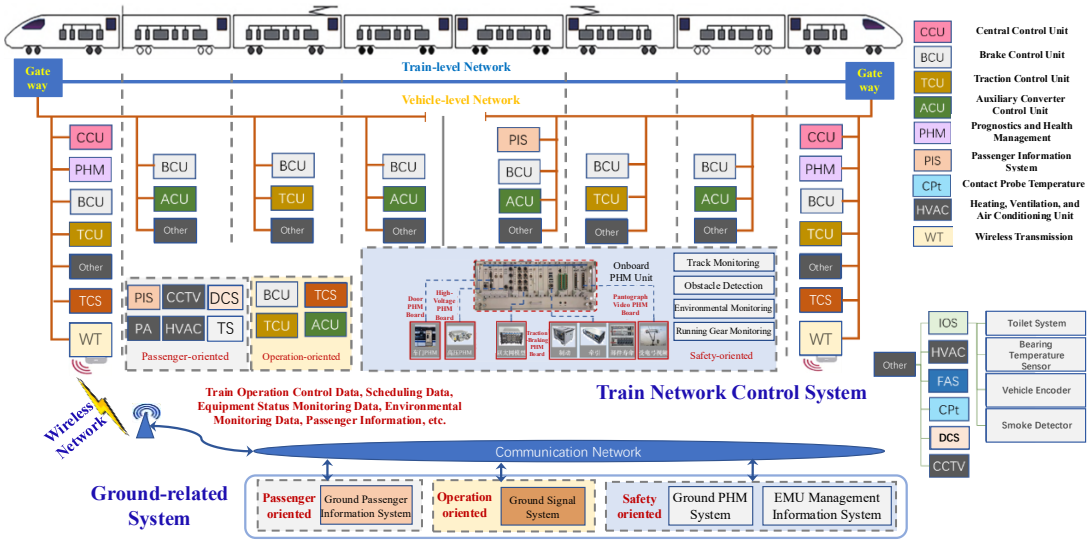

Figure 1 The Train Network Control System and Its Interfaces

1.1 Analysis of Current Status of Train Network Control

The train network control system performs unified control, monitoring, and basic diagnosis. However, further improvements in train safety, reliability, and operational efficiency are challenging due to the limited digitization of onboard devices, lack of autonomous train operation, and insufficient intelligent decision-making. The core challenges may be addressed by using comprehensive spatiotemporal perception of the operating environment, real-time dynamic monitoring of train equipment status, and data-driven operational decision-making and control. Nevertheless, due to the integration of third-party systems (e.g., the Passenger Information System (PIS) interfacing with ticketing systems) and the semi-open nature of train-ground communication networks, the train network faces critical challenges such as low data security, slow information transmission, and ineffective data integration and utilization. These challenges include the following aspects:

(1) Low Data Security Due to Reliance on Traditional Passive Protection Measures

The train network control system primarily relies on conventional boundary-based security mechanisms like firewalls. However, embedded units with limited computational resources and cross-platform characteristics may fail to support advanced security features. Additionally, inevitable vulnerabilities in operating systems expose the network to external attacks in non-fully enclosed environments, with complicated situation of safeguarding the train network data.

(2) Privacy Disparities Impeding Secure Data Sharing Across Multi-Vendor Equipment

During the train operation, computational units often experience insufficient processing limitations. For instance, the onboard Prognostics and Health Management (PHM) unit of the Fuxing Hao Train monitors over 3,000 data points while simultaneously performing real-time fault diagnosis and predictive alerts, occasionally encountering computational bottlenecks. Thus, collaborative support from idle computational units is necessary. But differences in privacy policies among equipments from different vendors and security levels, coupled with inadequate secure sharing and privacy protection mechanisms, hinder the cross-unit data sharing and collaborative resource utilization, with decreased task execution efficiency.

(3) Challenges in Efficient and Trusted Transmission of Sensitive Data in Semi-Open Communication Environments

Autonomous train operation and health management rely heavily on the accurate collection and efficient transmission of sensitive data, such as environmental

perception data and equipment status information. However, the train-ground communication depends on hybrid public-private networks: public networks offer high bandwidth but suffer from lower reliability, while private networks prioritize security at the expense of transmission speed. Consequently, the balance between transmission efficiency and security in semi-open environments remains a significant challenge.

On the other hand, considering the heterogeneous nature of train network resources, the sensitivity of inter-resource interactions, and the semi-open nature of communication networks, constructing a decentralized data security framework and a privacy-preserving computing architecture for high-speed railway networks raises the following challenges:

(1) Comprehensive Distributed Security Protection for Train Network Data in Heterogeneous Resource Scenarios

Autonomous train operation and equipment health management rely on the collaboration of distributed computing and storage resources within the train network control system. However, system vulnerabilities will be introduced due to hardware diversity, including variations in device origin, model, and specifications coupled with software heterogeneity across multiple vendors with different functionalities, performance levels, and cross-platform operations. These factors increase the complexity of data backup, access control, and identity authentication, bring challenges to the establishment of a unified security framework.

(2) Ensuring Secure Data Sharing in Train Coupling Scenarios

Collaborative resource utilization in the train network control system inherently depends on data sharing and integration across different units. However, frequent coupling and decoupling of different types of trains with varying configurations lead to dynamic network topologies and heterogeneous data structures, thus bring complicated centralized management. Additionally, privacy disparities among multi-vendor units, along with varying privacy policies, severely hinder secure data sharing and integration. Furthermore, the increasing number of wireless access points in coupled trains raises the risk of external intrusions, amplifying threats of data tampering and leakage. Therefore, secure cross-unit data sharing in such dynamic environments remains a critical challenge.

(3) Real-Time Secure Interaction of Sensitive Data in Semi-Open Communication Environments

In the train network control system, the interactions between onboard units and train-ground systems involve high-concurrency transmission of safety-sensitive data, such as control commands, equipment status, and PIS-related ticketing data. However, the hybrid public-private wireless communication networks where public networks offer high speed but low security, and private networks prioritize security at the cost of bandwidth, make it difficult to achieve the optimal balance between transmission efficiency and data protection.

It is seen from above analysis that to enhance autonomous operation and intelligent health management in high-speed railways, it is vital to develop data security mechanisms for shared and collaborative processing in train network control systems. Future issues may include decentralized storage features, as well as advanced technologies such as trusted computing, blockchain, and federated learning.

Scientific Problem Statement

Although numerous studies have explored data privacy and security in intelligent railway systems, most existing approaches either rely on passive protection mechanisms or lack compatibility with dynamic coupling/decoupling scenarios. Moreover, evaluation frameworks for current methods often overlook real-time performance, cross-unit interoperability, or resilience to emerging threats (e.g., adversarial attacks or quantum computing risks).

For instance, traditional cryptographic schemes cannot be seamlessly deployed on low-power onboard devices, and centralized architectures face bottlenecks in data integration and decision-making. Furthermore, existing assessments rarely consider the unique structural and operational constraints of high-speed train environments.

These limitations form the core problem that this paper seeks to address, i.e., how to design a scalable, secure, and privacy-preserving data sharing and processing architecture for semi-open, heterogeneous, and dynamic train networks.

1.2 Research Objectives and Contributions

Despite the existing studies on privacy and security, the applicability in railway environments is frequently ignored. To address this issue, this paper reviews current decentralized data sharing methods and privacy-preserving techniques, and discusses their feasibility for application in railway contexts. The review of relevant studies on train data privacy protection is shown in Table 1.

Table 1 Review of relevant studies on train data privacy protection

References	Field	Contribution	Upcoming issues
Chen et al.[1]	Detection and Diagnosis of Faults	The data-driven methods proposed in this paper (e.g., edge computing and transfer learning) can reduce sensitive data exposure through local data processing and model sharing, thereby indirectly enhancing the privacy protection.	Core privacy-preserving technologies such as data encryption and anonymization are not explicitly addressed, and the privacy risks associated with centralized data storage and transmission remain inadequately discussed.
Sun et al.[2]	Train Control Network	The paper advocates employing end-to-end encryption, rigorous access control, and integrated intrusion detection to protect the high-speed train operational data, while ensuring standardization and intelligent defense.	Current studies focus on isolated technical validations, without considering a universal security framework, real-time monitoring, and empirical support mechanism.
Kour et al.[3]	Railway Infrastructure	The paper proposes a blockchain-based "edge-cloud" collaborative architecture combined with lightweight encryption schemes to achieve data integrity protection and privacy enhancement. Additionally, blockchain technology is applied to maintain registration systems, ensuring trustworthy maintenance data.	Besides rolling stock and signaling systems, the energy and human factors should be involved, as well as the security assessment of third-party devices especially for LTE-R(Long Term Evolution - Railway)/5G(5^{th} Generation) networks.
Zubaydi et al.[4]	Internet of Things	The blockchain technology ensures data integrity and privacy through decentralized storage and smart contracts, combined with lightweight encryption schemes. Consortium blockchain architectures enhance the scalability and security, thus suitable for trustworthy sharing of multi-source high-speed train data.	Blockchain may be used considering high resource consumption and real-time data processing; Sufficient security assessments are needed for railway-specific communication protocols and unaddressed security risks in the third-party device integration.
López et al.[5]	Railway Transportation Systems	The article highlights blockchain and AI technologies for enhancing data privacy, with standardized frameworks and user awareness for secure railway systems. Enhanced LTE-R protocols and AI-driven anomaly detection improve the real-time data protection.	Detailed exploration are needed for data encryption specifics and emerging technologies (e.g., quantum encryption), with analysis of technical feasibility and cost-effectiveness in practical deployment, to address the security risks of third-party device integration.
Wang et al.[6]	Mobile Smart Devices	Privacy techniques in mobile crowd sensing, such as *k*-anonymity, homomorphic encryption, and differential privacy can protect identity, location, and content privacy of high-speed train data.	Existing solutions rely heavily on trusted third parties, vulnerable to single-point attacks; high computational/communication may affect real-time performance.
Yang et al.[7]	Cloud Storage	Attribute-based encryption (ABE) and searchable encryption enable fine-grained access control and secure retrieval for high-speed train data; integrity schemes can prevent tampering, while anonymous CP-ABE helps to enhance the privacy.	The fully homomorphic encryption brings higher computational costs for real-time processing; existing models assume semi-honest clouds, overlooking malicious adversaries; practical solutions are needed for real-time encryption and transmission efficiency in massive dynamic datasets.
Rafiq et al.[8]	Big Data Applications	Differential privacy and *k*-anonymity can enhance data anonymization, while hierarchical security frameworks can support systematic protection of multi-source data.	Optimization for real-time dynamic data streams are needed to improve the efficiency of encryption in high-frequency processing.
Khan et al.[9]	Future Transportation System	A novel paradigm for collaborative intelligent systems is introduced, featuring a standardized multi-layer lifecycle and optimized data management for federated learning. A secure modular framework based on distributed ledger technology (DLT) is designed to ensure transaction integrity across network nodes.	Current implementations of cloud, fog, and edge computing lack efficient coordination in transportation data management, often resulting in high latency and significant resource overhead.

Summary of Contributions

To address the identified challenges, this paper provides the following key contributions:

(1) We conduct a comprehensive and structured review of existing decentralized data sharing and privacy-preserving techniques, highlighting their strengths, limitations, and applicability to the railway domain.

(2) We design a blockchain-based asymmetric encryption framework tailored for secure data storage and cross-train communication in coupling/decoupling scenarios.

(3) We propose a federated learning-based collaborative computing framework that ensures privacy-preserving data sharing while addressing computational constraints in onboard environments.

(4) We identify and discuss future research directions, including quantum-resistant security architectures and adaptive privacy mechanisms suitable for intelligent railway systems.

The scope of this study is confined to train onboard data exchange and collaborative modeling in coupling/decoupling scenarios. It does not cover trackside infrastructure security, emergency communication protocols, or energy systems.

The rest of this paper is organized as follows:

Section 2 conducts a layered analysis of emerging technologies potentially applicable to current railway transportation.

Section 3 discusses existing decentralized technologies and their application scenarios.

Section 4 analyzes some typical privacy-preserving techniques.

Section 5 summarizes the findings and proposes a blockchain-based encrypted storage and collaborative computing architecture.

Section 6 concludes the paper and outlines future research directions.

2 Review of Data Security and Privacy Protection Methods in Railway Transportation

2.1 Hardware Layer

The train communication network (TCN) serves as a fundamental component of railway communication systems, interconnecting onboard equipment to facilitate efficient data exchange and ensuring the coordinated operation of train control systems. In June 1996, the establishment of IEC 61375 and UIC 556 as international standards set a foundation for intra- and inter-vehicle communication, promoting the interoperability among vehicles from different manufacturers and fostering the development of the railway industry[10]. However, railway systems rely heavily on onboard and trackside sensor networks, which leaves the systems open to a variety of security threats. By reviewing standards and evaluating railway cyberattack cases, this paper identifies vulnerabilities in current communication protocols and proposes a hierarchical security framework for railways. This framework integrates multiple methodologies to enhance the network resilience[11]. With technological advancements, the communication burden on trains has grown significantly. While facing external security threats, the internal communication design, which has aged over time, requires updating to accommodate escalating data demands. The fog radio access network is considered as a promising solution, but its high-performance requirements limit its applicability in complex railway communication scenarios. Some studies use dynamic network power allocation methods to reduce total network costs. For example, optimizing instantaneous power distribution of remote radio heads under multiple Quality of Service (QoS) constraints achieves reductions in network power [12].

Embedded systems onboard trains now require stronger security to prevent program failures or privacy breaches under cyberattacks. The Security Monitoring Unit (SMU)-integrated embedded system, as a System-on-Chip(SoC) module, ensures secure program execution and data processing, detects unauthorized instruction modifications and three types of data tampering attacks, and maintains low performance cost, balancing security and resource efficiency[13]. Beyond train control systems, edge devices such as door control units also face security threats. A hardware root-of-trust architecture tailored for low-power edge devices was proposed, featuring an accelerator-based SoC design. This architecture protects the execution environment integrity in uncontrolled deployment scenarios by isolating application software and safety-critical software states through access policies, achieving robust protection with minimal hardware cost [14]. To counter emerging quantum computing threats, traditional asymmetric encryption methods require upgrades. A post-quantum secure boot solution was proposed and fully implemented in hardware. This research employs the Extended Merkle Signature Scheme, a hash-based method, which demonstrates competitive performance compared to fully hardware-implemented Elliptic Curve Digital Signature Algorithm solutions[15]. The Train Control and Monitoring System enhances the train efficiency and safety to some

extent. However, as railway systems increasingly rely on automation, control, and communication technologies, networked control systems expose them to more cyber-physical security threats. Some studies analyzed system vulnerabilities, discussing the direct impacts of various attacks on functionalities and potential cascading consequences. Specifically, a systematic security risk assessment methodology was proposed, providing a reference for future railway security measures[16].

2.2 Software Layer

Software security challenges in railway communication networks arise from multiple risk factors, including cross-platform compatibility, operating system vulnerabilities, software updates and maintenance, software design, and data privacy.

TCN rely on message buses, which makes it challenging to support distributed monitoring and intelligent fault diagnosis in complex environments. A study proposed a layered ontology model integrating Ethernet/IP protocols and common object request broker architecture middleware to construct a three-tier architecture (subsystem, carriage, and train levels). This approach achieves structured representation and reasoning of domain knowledge through semantic modeling, offering mechanisms for lifecycle management, service discovery, and information aggregation to enhance fault diagnosis and maintenance efficiency [17].Similarly, another study introduced a flexible hierarchical architecture leveraging Software-Defined Networking and Network Function Virtualization for dynamic network orchestration, Multi-access Edge Computing to reduce latency for critical services, blockchain to strengthen data security, and AIto optimize resource scheduling and fault prediction. Some studies further proposed end-to-end network slicing schemes tailored for Ultra-Reliable Low Latency Communication (URLLC), evolved Mobile Broadband, and massive Machine-Type Communication (mMTC)[18].

With increasing demands for high availability, scalability, and efficient hardware resource utilization in railways, migrating safety-critical applications from hardware platforms to cloud platforms presents a viable solution. After comparing two widely used virtualization technologies, KVM and Xen, the study analyzed their suitability for meeting stringent security and real-time requirements in railway systems. The RT-Cloud framework, based on commercial off-the-shelf hardware, enables resource sharing via virtualization to reduce costs and improve flexibility. Additionally, a novel resource management layer was proposed to ensure effective resource allocation for real-time safety-critical applications[19]. However, the complex structure and ambiguous boundaries of railway communication networks introduce security threats like information leakage and malicious access when cloud computing is adopted. To address this, a zero-trust security model was proposed, utilizing blockchain and Merkle trees to construct a distributed identity storage scheme. The model incorporates proxies for mutual authentication with cloud servers, enhancing system security, efficiency, and stability while preventing malicious external devices from compromising safety[20].

2.3 Network Layer

The traditional Global System for Mobile Communica-tions – Railway(GSM-R) system, which is based on 2G technology, suffers from inherent weaknesses such as outdated protocols and weak encryption. In contrast, emerging 5G/WiFi-integrated architectures enhance the transmission efficiency but introduce new attack risks. A study proposed a multidimensional threat analysis framework[21]. This framework incorporates advanced encryption protocols, multi-factor authentication, and network slicing as mitigation strategies. Through threat modeling and component-level security enhancements, this framework significantly improves the anti-jamming capabilities of railway communication networks**.** To address similar opportunities and challenges, another study adopted 5G technology as pivotal for future railway mobile communication systems. Key 5G features like flexible subcarrier spacing, massive MIMO, network slicing, URLLC, and mMTC, provide solid support for the next generation of railway industry[22].

External threats like signal jamming attacks also affect the train operational safety. A frequency hopping spread spectrum-based jamming mitigation method was investigated, demonstrating promising results in simulated environments[23] . To address the lack of decentralized authentication in existing communication protocols, a blockchain was integrated to enhance the Communication-Based Train Control (CBTC) network security[24]. Combined with partially observable markov decision process, this approach derives an optimal adaptive consensus strategy balancing network security and efficiency, reducing the impact of data tampering attacks on train operations.

Similarly, with the leveraging spread spectrum technology, a study proposed direct sequence spread spectrum using a cryptographically secure pseudo-random number generator [25]. This method transforms an attacker's jamming signal into Gaussian noise post-reception, effectively suppressing the interference. For Sybil attacks, a novel CBTC system was designed, integrating local security authentication and collaborative

security checks with asynchronous reinforcement learning based on the quantized age of information[26]. This system achieves higher probabilities of Sybil attack detection and defense.

For growing service demands and limited broadband resources in train environments, a RAN slicing-based wireless train communication network was introduced[27]. By employing joint bandwidth optimization and terminal clustering algorithms, this architecture efficiently allocates slice bandwidth to train control services, passenger information services, and train sensing services, improving system performance.

Another study integrated AI with classical optimization techniques, demonstrating the effectiveness of deep learning and game theory in joint optimization. For instance, federated learning was utilized to balance the bandwidth and training latency while preserving privacy[28]. Specifically, for millimeter-wave train-ground communication systems, researchers integrated the full-duplex technology with onboard mobile relays to enhance the system capacity and performance. Thus, a low-complexity transmission scheduling algorithm was designed[29].

The data security and privacy protection of railway systems require multi-level collaborative innovation. At the hardware level, security monitoring units and post-quantum cryptographic mechanisms enhance protection capabilities. However, scalable solutions are urgently needed to address aging infrastructure. At the software level, the integration of a zero-trust architecture and intelligent resource management aims to balance cloud adoption needs, but the real-time performance remains a challenge. At the network level, the adoption of 5G integration, anti-interference technologies, and network slicing improves the communication reliability but should meantime ensure the compatibility with legacy protocols. Future efforts should focus on post-quantum cryptographic systems, standardized security evaluation frameworks, and adaptive coordination mechanisms to address systemic risks posed by quantum computing, cross-domain attacks, and heterogeneous network integration. These efforts will contribute to building a resilient and efficient intelligent railway security ecosystem.

3 Review of Decentralized Data Sharing Methods

Decentralized systems can effectively address privacy leakage issues caused by single points of failure. This section reviews existing decentralized data security sharing technologies and their applications.

3.1 Distributed Technologies

Distributed technologies improve the availability, scalability, fault tolerance, and flexibility of train network control systems. These benefits are achieved by distributing tasks and data across multiple computational nodes. This approach distributes system components across nodes, enabling collaborative operation in distributed environments[30]. Azad et al.[31] designed a hierarchical, parallel, and scalable distributed I/O system for distributed database storage. This system integrates GPU kernel acceleration with a combined BLAS library to enable rapid algorithm development. To address data integrity verification, Zhang et al.[32] proposed a multi-file, multi-replica batch auditing scheme. This approach reduces storage cost while enhancing security. The performance of metadata services significantly impacts the overall distributed system performance[33]. Gao et al.[34] introduced the first machine learning-based DeepHash model, which preserves metadata locality and balances loads across metadata servers. Zhang et al.[35] proposed a novel solution designated as SMURF to enable the distributed continuum caching and semantic locality-aware prefetching strategy, thereby achieving reliable low-latency transmission. Rapp et al.[36] proposed the DISTREAL, a resource-aware adaptive learning mechanism for distributed training. This approach targets heterogeneous and resource-constrained environments within machine learning-driven distributed computing engines. Wang et al.[37] developed a robust distributed deep forest framework using a two-stage pre-aggregation method to adjust class vector granularity, accelerating task recovery with minimal system resources. For intrusion detection, Liu et al.[38] designed a hierarchical distributed intrusion detector tailored to industrial cyber-physical systems, providing comprehensive security protection by aligning with layer-specific system architectures and attack patterns.

3.2 Blockchain Technology

To mitigate the limited encryption capabilities of distributed technologies, blockchain enhances data security through cryptographic techniques and consensus algorithms, ensuring transaction transparency and immutability. Wu et al.[39] proposed a blockchain-based attribute proxy re-encryption data sharing strategy. This strategy dynamically updates distributed key generation methods to mitigate key leakage risks. To enhance encryption/decryption security, Shalini et al.[40] replaced traditional key distribution with quantum key distribution .Zhang et al.[41] developed a privacy-preserving data security sharing model for blockchain-driven Industrial Internet of Things (IIoT) environments. Khan et al.[42]

proposed a novel architecture, BDLT-IoMT, which integrates blockchain distributed ledger technology with Support Vector Machine (SVM)-based machine learning. By embedding SVM algorithms into the blockchain framework, the architecture enhances data classification, resource scheduling, and node communication efficiency, thereby addressing computation overhead and scalability challenges in distributed environments. In industrial collaboration, Guo et al.[43] designed a hybrid concurrency control protocol on a blockchain architecture, resolving consistency and concurrency challenges in heterogeneous data sharing. Yuan et al.[44] introduced CoopEdge+, a blockchain-based decentralized platform. It addresses collaborative computation challenges in scenarios involving untrusted edge servers.

Recent studies have explored the integration of blockchain with trusted hardware to enhance privacy and performance. Wang et al.[45] proposed a novel architecture named HybridChain, which combines blockchain with trusted execution environments (TEEs) and decouples computation from consensus via a hierarchical network structure. This design improves both the confidentiality and performance limitations of traditional blockchain systems. Similarly, Liu et al.[46] addressed the mismatch between on-chain and off-chain environments by designing a cost-efficient mapping protocol based on TEEs. Their work includes a detailed analysis of attack resilience and provides valuable guidance for future secure implementations.

In another line of research, hardware security modules (HSMs) have been utilized to ensure robustness and trustworthiness in blockchain-based applications. Castillo et al.[47] proposed an efficient integration scheme of HSMs with public key cryptographic algorithms and blockchain systems, thereby establishing a trusted root for all monitored data extraction processes.

To enhance encrypted search capabilities, Guo et al.[48] incorporated Dynamic Searchable Symmetric Encryption (DSSE) into blockchain systems to support forward-private encrypted queries. They further introduced a hybrid indexing mechanism to offset performance overhead induced by DSSE. In parallel, Linoy et al.[49] developed a distributed computing platform supporting smart contracts, where queries are transformed into MapReduce tasks executed on Hadoop to improve computational efficiency. Security is preserved through encryption and proxy-based mechanisms. Khan et al.[50] designed a lightweight consensus mechanism based on advanced practical Byzantine Fault Tolerance, tailored for resource-constrained IoMT scenarios. By combining edge computing with hybrid encryption techniques, the framework enables lightweight authentication suitable for low-power devices, effectively overcoming the limitations of conventional blockchain systems in such environments.

3.3 Federated Learning Technology

Federated Learning (FL), emerging as a solution to data privacy concerns, seeks to address the issue of data warehouse. However, risks such as model inversion and model extraction attacks continue to pose threats of data leakage. Researchers have integrated privacy and secure computation methods into FL. Guo et al.[51] applied a two-tier differential privacy mechanism to protect privacy in cloud-edge-device training models. Huang et al.[52] addressed the trade-off between privacy protection and cost in FL with differential privacy by proposing a Stackelberg game-based framework, incentivizing client-server collaboration through reward mechanisms. Wang et al.[53] introduced a differential privacy strategy based on non-Gaussian noise. This method encrypts/decrypts local features and complicates statistical inferences, thereby deterring attackers. However, these methods belong to passive defenses. Facing system-level challenges like device heterogeneity, most synchronous FL paradigms suffer inefficiency due to the straggler effect. Li et al.[54] proposed a theoretically driven multi-stage adaptive private algorithm to balance the model utility and privacy in differentially private asynchronous FL. Homomorphic encryption (HE), superior in complexity to differential privacy, integrates privacy protection into FL. Xu et al. [55] applied HE to FL and smart contracts, resolving privacy and trust issues in industrial IoT. Jing et al.[56] enhanced a multi-key homomorphic encryption protocol, encrypting model updates via aggregated public keys before server-side aggregation, with improved accuracy and cost efficiency.

While distributed technologies eliminate single-point privacy leaks, the blockchain ensures data security through immutability and encryption. Besides, the FL enables model training without data sharing, the semi-open nature of train network control systems and encryption performance limitations hinder the standardization and interoperability across systems, thus restricting the data sharing efficiency and scope. These issues deserve further research.

4 Review of Privacy-Preserving Methods

Secure data sharing technologies address computational resource shortages, protect the sensitive information, and ensure the data integrity. However, distrust among data owners, coupled with leakage risks and conflicting interests, increases maintenance costs and data

risks despite of growing sharing demands.

4.1 Privacy-Preserving Computational Methods

Existing privacy-preserving methods are frequently integrated with FL. For instance, differential privacy is commonly used to protect data by introducing noise to raw data or model outputs. He et al.[57] applied the local differential privacy to train clustered FL models. This approach perform model training with heterogeneous IoT data while reducing the noise and communication cost. Ren et al.[58] proposed a secure distributed stochastic approximation method combining FL and differential privacy, optimizing the power system performance while preserving the privacy. Wei et al.[59] designed a privacy-preserving FL scheme for deep learning. By integrating the secure multi-party computation (SMPC) with differential privacy, this method prevents data theft and mitigates risks of malicious inference from shared global information. Despite these benefits, computational resource constraints and algorithmic complexity in train network control systems due to data volume and hierarchical complexity still limit the response speed, accuracy, and scalability.

In railway applications, decentralized data sharing and privacy methods are widely adopted. For instance, Zhang et al.[60] combined distributed optimal control algorithms with virtual platooning for real-time train unit control. Zhang et al.[61] proposed a blockchain-based cloud key management system for high-speed railways, enhancing communication security and reducing latency. Liang et al.[62] leveraged Edge Intelligence to provide real-time services for Urban Rail Transit , designing a blockchain-based trust management mechanism to improve learning efficiency and resource utilization. For fault diagnosis, Qin et al.[63] proposed the enhanced FL using wavelet packet decomposition to reduce the computational costs and using SecureBoost for training local diagnostic models. For traffic data sharing, Jiang et al.[64] employed the blockchain for multi-party secure collaboration, addressing data flow barriers in railways.

Although decentralized systems eliminate single-point failures, use encryption to prevent leaks, reduce latency, and enhance reliability, existing methods still show shortcomings like:

- Complex data protection levels and device heterogeneity in train networks.
- Uncontrolled computational costs and performance of edge devices.
- Limited focus on multi-network interactions in railway systems.

4.2 Sensitive Data Transmission Methods

Lim et al.[65] investigated threats to data integrity in balise transmission modules. Their study focused on challenges arising from the informatization of train-ground communication. Using high-fidelity simulations to study risks from data integrity attacks, a secure hybrid train speed controller was designed, though with limited efficiency. Vahidi et al.[66] proposed a transmitter-receiver architecture for data transmission and detection in 6G communication systems between high-speed trains and base stations. To address emerging demands for train network informatization, Wang et al.[67] adapted and refined a deep convolutional neural network model (AlexNet). By enhancing network layers and learning capabilities, their approach achieved high data transmission security without compromising prediction accuracy. For fault diagnosis, Du et al.[68] introduced a federated transfer learning framework for fault diagnosis. This framework provides an effective solution for zero-shot fault diagnosis in high-speed train bogies. Saki et al.[69] proposed three algorithms for optimal access point placement to enhance wireless train-ground communication, thus improving the data transmission reliability. To meet the increasing requirements for reliability and latency in railway mobile data communication, Wang et al.[70] designed a parallel redundancy protocol for railway wireless networks, ensuring robust data exchange between onboard and trackside devices. While novel wireless systems enhance the transmission speed, stability, and security, their design complexity and high costs often limit the functional realization.

In railway data security frameworks, Chan et al.[71] proposed a security deployment framework as a reference for cybersecurity testing. Luo et al.[72] improved data transmission timeliness, reliability, and security by leveraging data mining and digital mobile communication, tailored to railway operational and train-control data characteristics. Soderi et al.[73] developed a cybersecurity assessment procedure and network range-based method to simulate and validate railway network system security. For railway environmental monitoring, Wang et al.[74] introduced an energy-balanced data transmission strategy for linear wireless sensor networks, addressing real-time, energy-efficient, and robustness demands. Wu et al.[75] designed a fuzzy algorithm-based 3D laser scanning data transmission system, significantly reducing the packet loss rates.

It is noted that existing methods rely on encryption algorithms and intrusion detection systems to partially ensure the security of sensitive data. Meanwhile, novel wireless systems and data exchange technologies have been developed to enhance the transmission efficiency. However, due to the inherent limitations and semi-open nature of train-ground communication networks, balancing security and efficiency remains challenging.

Thus, achieving efficient and trustworthy transmission of interactive data between train-ground and intra-train network nodes remains a critical area for systematic research.

5 Blockchain-Based Asymmetric Encrypted Storage and Collaborative Privacy-Preserving Computing Architecture

5.1 Blockchain-Based Data Security Storage Mechanism

To address the traceability and tamper-proof requirements of train network data, we propose a blockchain-based secure storage mechanism. First, considering the massive data and privacy constraints in train coupling scenarios, a consortium blockchain-based network data security architecture is designed. Building upon this architecture, a Raft consensus algorithm tailored for train network consortium chains is developed to enhance consensus efficiency and support secure, high-speed data storage. A Merkle tree based system data integrity verification method is employed to overcome the limitations of traditional data integrity verification frameworks in handling dynamic data, thus ensuring the secure and efficient storage.

Given the traceability and anti-tampering needs of onboard PHM units and wireless transmission units, the proposed blockchain-based data security storage architecture is illustrated in Fig 2. To meet the demands for large-scale data storage and high-speed processing in PHM and wireless units, the efficient Raft consensus algorithm is adopted. The Merkle tree based verification method safeguards the data security and integrity.

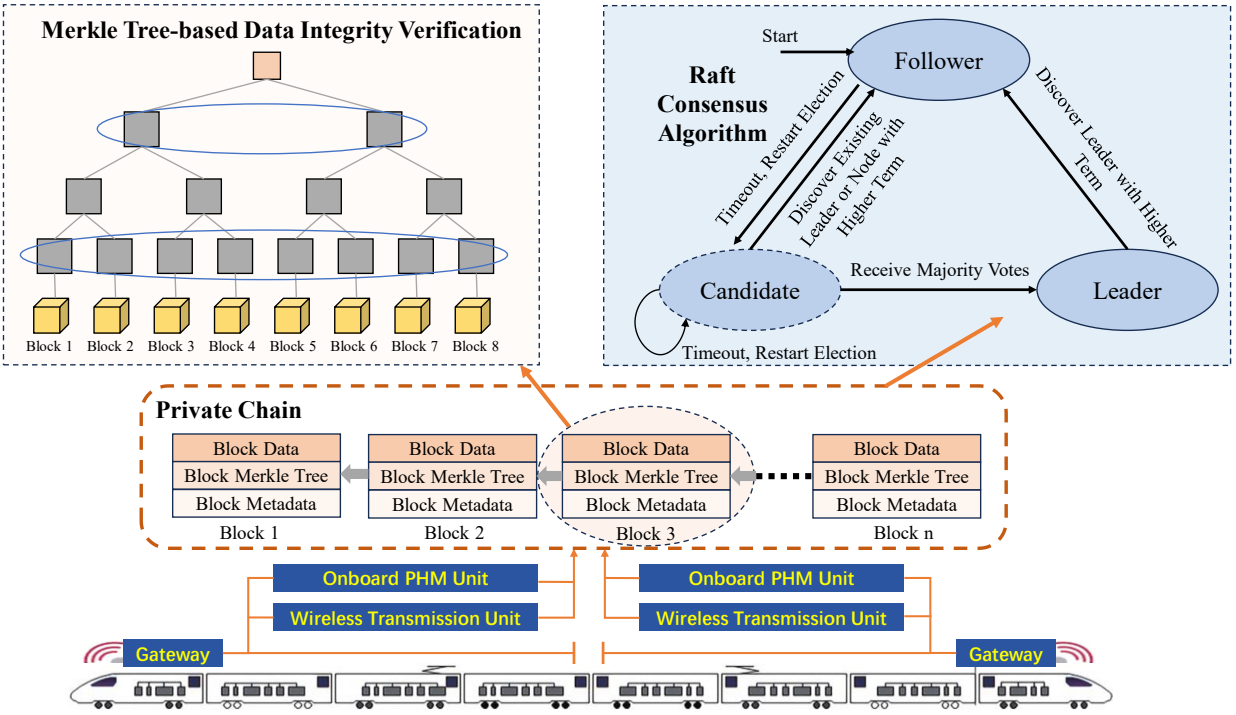

Figure 2 A suggested Secure Data Storage Model based on Blockchain

5.1.1 Consortium Blockchain-Based Train Network Data Security Architecture

First, to address the challenges of massive data storage and cross-chain sharing in train networks, a hybrid on-chain/off-chain collaborative architecture is designed. Since the operational data (e.g., equipment status, fault diagnostics) generated during train operation is enormous, directly storing all data on-chain would severely decrease the blockchain efficiency or even lead to system crashes. Thus, an off-chain distributed storage scheme is implemented: the raw data is categorized and stored locally on onboard cards and PHM units, while the blockchain records only critical metadata (e.g., data identifiers, timestamps) and the Merkle tree root hashes for integrity verification. This design preserves tamper-

proof nature of the blockchain while avoiding the overload of on-chain storage.

In train coupling scenarios, a single train's private chain must interact with other train chains. Heterogeneous data structures or divergent consensus mechanisms across private chains bring compatibility challenges. To resolve this issue, cross-chain smart contracts are deployed to build a consortium blockchain framework:

- Unified data exchange format standards are defined to ensure semantic consistency across chains.
- Cross-chain communication protocols are established, specifying identity authentication rules and encrypted transmission mechanisms.

For example, when two coupled trains need to share bearing temperature data, the initiating chain triggers a cross-chain request via smart contracts. The target chain validates the request's legitimacy, verifies data integrity via Merkle trees, and completes data synchronization through encrypted channels.

The core module, i.e., the cross-chain smart contract, integrates multi-layered security controls. It employs a dynamic permission model to restrict access to data of varying security levels (e.g., general operational parameters vs. critical control commands) and utilizes timestamp and hash verification to prevent replay attacks. The consortium blockchain adopts a dynamic leader election mechanism, automatically selecting primary nodes based on real-time computational power and data interaction frequency, ensuring cross-chain collaboration stability during network fluctuations or node failures. This design achieves efficient and secure collaboration among coupled trains while preserving data traceability and privacy.

5.1.2 Raft Consensus Algorithm for Train Network Consortium Chains

To meet the high-frequency on-chain storage demands of PHM units and wireless transmission units in train networks, the Raft algorithm is adopted to achieve efficient distributed consistency. Compared to traditional Practical Byzantine Fault Tolerance (PBFT) algorithms, the Raft significantly reduces consensus complexity through simplified communication mechanisms, making it suitable for resource-constrained onboard environments. The leader node batches and pushes log entries to followers, completing data submission after confirmation by a majority of nodes, thereby avoiding network congestion caused by full-node broadcasts. The key mechanisms include:

- Dynamic Leader Election: Random election timeouts (150-300 ms) enable rapid response to node failures, e.g., electing a new leader within 10 seconds if a PHM unit disconnects.
- Log Replication Optimization: Incremental replication strategy transmits only differential log entries.
- Byzantine Fault Resistance: Digital signatures verify log sources to counter malicious node data forgery.

5.1.3 Merkle Tree Based Data Integrity Verification Method

To verify the dynamic train data efficiently, a lightweight Merkle tree verification framework is designed. Specifically, data streams collected by PHM units are segmented into dynamically adjustable blocks based on computational capacity. The hash values for each block are computed to construct a four-layer Merkle tree. While the root hash is stored on-chain, the complete tree structure is retained locally for rapid verification. For example, when verifying brake pressure data from a specific period, the corresponding root hash is retrieved from the blockchain, and 15% of data blocks are randomly sampled and rehashed. By comparing hash paths layer-by-layer in the Merkle tree, problematic blocks are precisely identified. Theoretically, this method significantly improves the verification efficiency over traditional CRC checks, with negligible false positives due to hash collisions.

5.2 Privacy-Preserving Data Sharing via Asymmetric Encryption and Smart Contracts

To solve the issues caused by sparse fault data from individual devices in coupled trains, this study proposes a collaborative solution integrating asymmetric encryption and blockchain smart contracts, enabling cross-unit trusted data sharing through a hierarchical security architecture.

5.2.1 Asymmetric Encryption-Driven Private Chain Data Sharing

Based on the train network topology (with two vehicle-level networks per train), PHM units and wireless transmission units are selected as nodes to construct private blockchains (four nodes per chain). The end-to-end secure transmission is achieved via the asymmetric encryption:

- Nodes generate RSA key pairs, with public keys shared on-chain.
- Senders encrypt data using the receiver's public key, ensuring that only authorized nodes can decrypt.
- Blockchain metadata tracing verifies data provenance and integrity.

This approach enhances the multi-train collaborative diagnostics in coupling scenarios and mitigates risks of sensitive data exposure.

5.2.2 Cross-Chain Smart Contract-Based Data Management

For secure data sharing across multiple chains in coupled trains, a hierarchical security governance architecture (Fig 3) is proposed, comprising four dimensions:

1. **Heterogeneous Chain Interoperability Protocol Layer**

As the foundational layer, its modular communication protocols enable trusted interconnections between private chains. Asymmetric encryption tunnels establish cross-chain channels, ensuring transmission security aligns dynamically with source-chain policies. This resolves interoperability challenges in data formats and consensus mechanisms.

2. **Multi-Dimensional Security Control Layer**

This layer integrates dual protection mechanisms:

- Dynamic Identity Authentication: A fine-grained permission matrix controls data access levels.
- Four-Tier Data Protection: To Classify data by sensitivity (operational parameters, fault features, device identifiers, control commands) for gradient security.

3. **Smart Contract Execution Engine Layer**

This layer adopts verifiable cross-chain routing contract clusters with core functions:

- Data routing selection.
- Integrity checks (e.g., timestamp anti-replay mechanisms).

State changes during execution are synchronized across participating chains in real time.

4. **Audit and Traceability Layer**

This layer establishes a blockchain-based credible audit system to record full interaction metadata. Standardized traceability APIs enable lifecycle queries for any data item, with immutability guaranteed by BFT consensus.

By vertically integrating protocol-control-execution-audit layers, this architecture constructs a full-cycle defense system, effectively tackling replay attacks, man-in-the-middle attacks, and other threats while maintaining data security levels and ensuring transparent, controllable sharing.

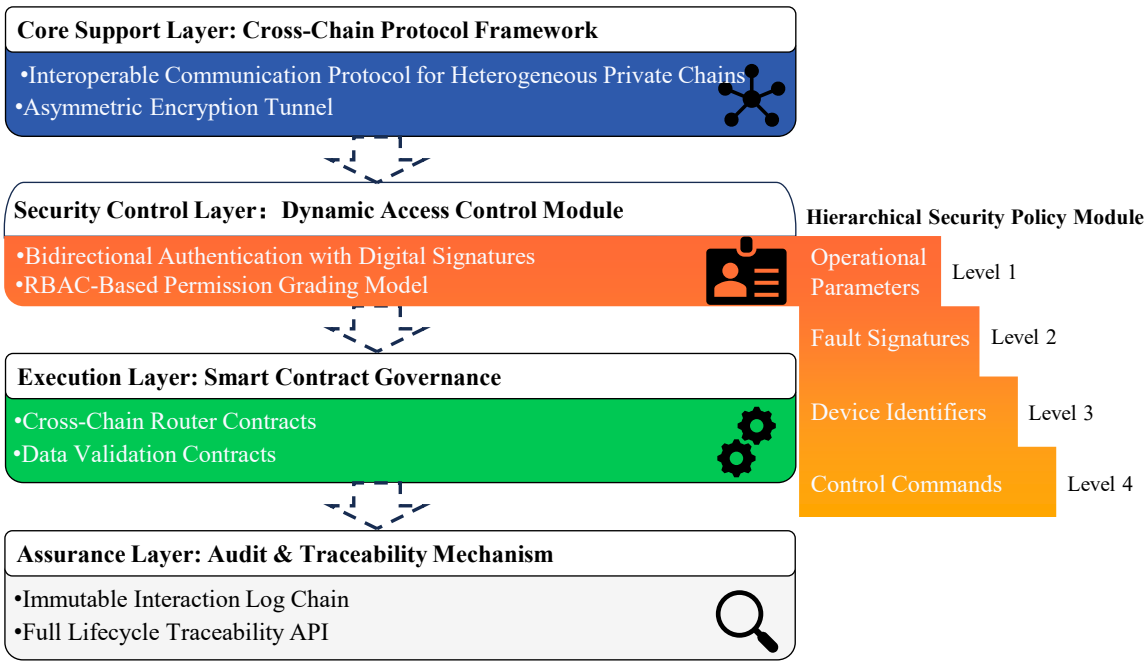

Figure 3 A Cross-chain Data Security Architecture Suggested for Upcoming Chinese High-speed Trains

5.3 Privacy-Preserving Computing-Based Secure Data Sharing and Collaborative Processing for Train Networks

To address the challenges posed by high-volume fault data, computational demands, and real-time requirements in onboard PHM units, this section proposes a privacy-preserving computing framework for secure data sharing and collaborative processing. By optimizing the

computational resource scheduling and leveraging idle units to supplement underpowered PHM units, this framework ensures normal operation while maintaining "usable but invisible" data privacy across heterogeneous security levels.

5.3.1 Container Virtualization Based Resource Optimization Scheduling

To resolve sudden computational demands in PHM units, a hierarchical resource scheduling strategy is implemented:

(1) Priority Allocation: High-security PHM boards are prioritized for processing sensitive core data.
(2) Dynamic Expansion via Container Virtualization:
- Real-Time Monitoring: Track container resource states such as power, memory, network load, and security levels.
- Resource Screening: Identify idle/low-load containers compliant with security policies.
- Dynamic Adjustment: Optimize the task allocation by adjusting CPU/memory quotas.
- Continuous Optimization: Enable dynamic container activation/deactivation and task migration through performance monitoring.

This scheduling strategy can ensure the data security and elastic computational resource scaling as sensitive data is only processed in high-security containers. By balancing security policies and load, a closed-loop dynamic resource management system is formulated, with enhanced system responsiveness to vibrating demand and maximized resource efficiency.

5.3.2 Federated Learning-Based Collaborative Data Processing

To address the dual challenges of insufficient computational power and data privacy protection of onboard PHM units, a collaborative computing framework is suggested which integrates the FL and data processing. This method achieves secure collaborative processing of high-security-level data through a distributed machine learning architecture, effectively resolving the conflict in traditional centralized computing models between sensitive data exposure risks and limited computational scalability.

At the system architecture level (as shown in Fig 4), each onboard PHM unit acts as a participant in federated learning, with a localized model training mechanism. First, the raw data are encrypted locally. Next, models are independently trained based on the encrypted data, and only securely encrypted model parameter updates (e.g., gradients) are transmitted to the central aggregator. The central aggregator constructs a global model through aggregation strategies such as weighted averaging of parameters, and the optimized model is redistributed to all participating units for iterative training. This mechanism avoids the privacy leakage risks caused by cross-node transmission of sensitive data.

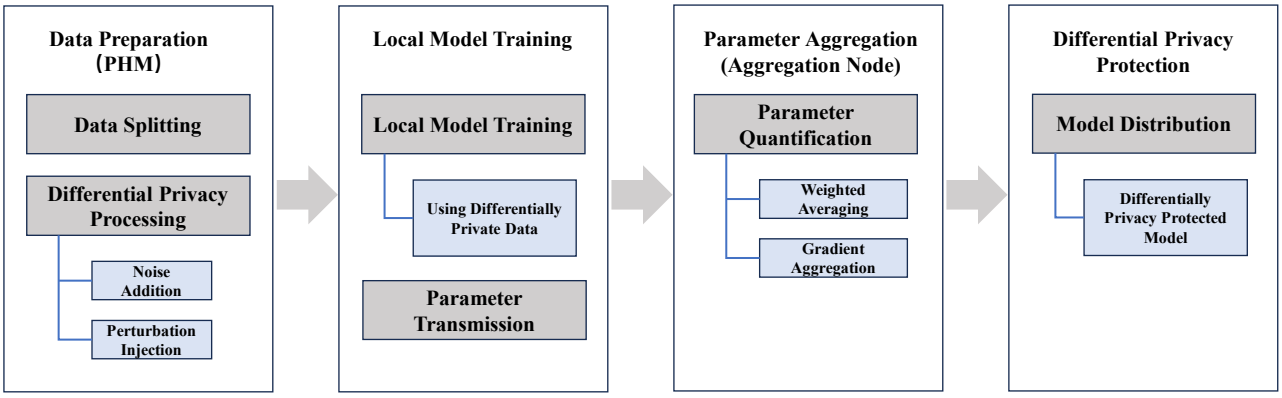

Figure 4 A Suggested Collaborative Data Processing Framework based on Federated Learning for Chinese High-speed Trains

To further enhance the privacy protection, differential privacy enhancement techniques are implemented throughout the federated learning lifecycle. In the local training phase, the noise perturbation is injected into model gradients, preventing the reverse engineering of raw data from parameter updating. In the model distribution phase, the global model undergoes privacy-preserving processing to mitigate the model inversion attack risks. With this dual-layer privacy protection architecture, the central server or malicious participants cannot identify individual data features accurately even if they can obtain the intermediate parameters.

- **Computational scalability:** The distributed computing framework effectively integrates idle computational resources in the vehicular network, avoiding the single-node computation.
- **Data security:** The dual privacy protection

mechanisms ensure high-security-level data processing ability.

- **Model efficiency:** The dynamic feedback mechanism optimizes the federated learning parameters like aggregation frequency and noise intensity, under privacy budget constraints, achieving a balance between model accuracy and convergence speed.

This solution provides an innovative paradigm for edge computing scenarios in vehicular environments. By performing optimization based on federated learning and differential privacy, a scheme is achieved to balance the privacy protection strength, computational scalability, and model performance, enabling more secure computing architectures in intelligent railway transportation systems.

6 Conclusions and Future Perspectives

This paper systematically reviews the data security and privacy protection challenges raised by modern high-speed railway transportation systems, especially in coupling/decoupling scenarios. We examine key advancements in decentralized storage, privacy-preserving computing, and trusted sharing mechanisms, analyzing security threats and mitigation strategies across hardware, software, and network layers.

- The study highlights the following main contributions: **Decentralized Data Management**: We explore how integrating blockchain with FL can enhance data traceability and privacy in railway systems. However, challenges remain in terms of interoperability among heterogeneous blockchain networks and the efficiency of consensus algorithms like PBFT in large-scale scenarios.
- **Evolution of Privacy-Preserving Technologies:** Techniques such as Differential Privacy and homomorphic encryption significantly improve sensitive data sharing security. However, challenges such as model accuracy degradation from noise injection and the need for stronger quantum-resilient methods still require further research.
- **Communication Architectures:** The integration of 5G and edge computing can improve train-ground data transmission with low latency and high reliability. However, issues related to data integrity and access control in semi-open networks remain unresolved, and the deployment costs of novel interference mitigation techniques need further evaluation.

Future Challenges Include:

- Heterogeneous system coordination, where divergent data formats and security levels across multi-vendor equipment decrease the cross-chain sharing efficiency.
- Privacy-computation trade-offs, especially in balancing encryption strength and real-time performance on resource-constrained edge devices.
- Slow dynamic threat response, with current intrusion detection systems relying on static rules and struggling to recognize emerging attacks like adversarial samples.

Potential Research issues:

- Quantum-resistant security architectures, focusing on the development of post-quantum cryptography methods for train networks and dynamic threat awareness mechanisms.
- Adaptive privacy-preserving frameworks, involving the integration of FL with Secure MPC to adjust privacy protection strength and model precision in real-time.
- Intelligent threat defense systems, leveraging federated reinforcement learning for distributed intrusion detection and collaborative adversarial attack recognition.

List of abbreviations

CCU	Central Control Unit
BCU	Brake Control Unit
TCU	Traction Control Unit
ACU	Auxiliary Converter Control Unit
PHM	Prognostics and Health Management
PIS	Passenger Information System
HVAC	Heating, Ventilation, and Air Conditioning Unit
WT	Wireless Transmission
LTE-R	Long Term Evolution - Railway
5G	5th Generation
ABE	Attribute-based Encryption
DLT	Distributed Ledger Technology
TCN	Train Communication Network
QoS	Quality of Service
SMU	Security Monitoring Unit
SoC	System-on-Chip
URLLC	Ultra-Reliable Low Latency Communication
mMTC	massive Machine-Type Communication
GSM-R	Global System for Mobile Communications – Railway
CBTC	Communication-Based Train Control
IIoT	Industrial Internet of Things
SVM	Support Vector Machine
TEEs	Trusted Execution Environments

HSMs	Hardware Security Modules
DSSE	Dynamic Searchable Symmetric Encryption
FL	Federated Learning
SMPC	Secure Multi-Party Computation
HE	Homomorphic Encryption
PBFT	Practical Byzantine Fault Tolerance

Author Contributions

Conceptualization, methodology, investigation, and writing—original draft preparation: Zhihang Zhang; Blockchain architecture design, project administration, funding acquisition, and writing—review and editing: Feng Wang; Federated learning design, supervision, research support, and writing—review and editing: Peng Li.

Figures Originality

All figures included in this manuscript are original and have been created by the authors specifically for this study. No figures have been reproduced or adapted from previously published materials. Additionally, no AI-generated tools or external software were used beyond standard academic diagramming software. Therefore, no copyright permissions are required.

Funding

This work is supported by the National Natural Science Foundation of China under Grant U2468203.

Conflicts of Interest

The authors declare no conflicts of interest regarding this manuscript.

Acknowledgments

The authors would like to acknowledge the assistance of AI tools, including ChatGPT, DeepSeek, and Doubao, for their support in language translation and proofreading during the preparation of this manuscript.

Reference

[1] Chen H, Jiang B, Ding S X, et al. Data-driven fault diagnosis for traction systems in high-speed trains: A survey, challenges, and perspectives[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 23(3): 1700-1716.

[2] Sun C, Zhang W, Wang H, et al. A Review of Research on the Security of Train Control Networks[C]//2024 6th International Conference on System Reliability and Safety Engineering (SRSE). IEEE, 2024: 459-463.

[3] Kour R, Patwardhan A, Thaduri A, et al. A review on cybersecurity in railways[J]. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 2023, 237(1): 3-20.

[4] Zubaydi H D, Varga P, Molnár S. Leveraging blockchain technology for ensuring security and privacy aspects in internet of things: A systematic literature review[J]. Sensors, 2023, 23(2): 788.

[5] López-Aguilar P, Batista E, Martínez-Ballesté A, et al. Information security and privacy in railway transportation: A systematic review[J]. Sensors, 2022, 22(20): 7698.

[6] Wang Y, Yan Z, Feng W, et al. Privacy protection in mobile crowd sensing: a survey[J]. World Wide Web, 2020, 23(1): 421-452.

[7] Yang P, Xiong N, Ren J. Data security and privacy protection for cloud storage: A survey[J]. IEEE Access, 2020, 8: 131723-131740.

[8] Rafiq F, Awan M J, Yasin A, et al. Privacy prevention of big data applications: A systematic literature review[J]. SAGE Open, 2022, 12(2): doi:21582440221096445.

[9] Khan A A, Yang J, Awan S A, et al. Artificial intelligence, internet of things, and blockchain empowering future vehicular developments: A comprehensive multi-hierarchical lifecycle review[J]. Human-centric Comput. Inf. Sci., 2025, 15: 13.

[10] Schifers C, Hans G. IEC 61375-1 and UIC 556-international standards for train communication[C]//VTC2000-Spring. 2000 IEEE 51st Vehicular Technology Conference Proceedings (Cat. No. 00CH37026). IEEE, 2000, 2: 1581-1585.

[11] Ibadah N, Benavente-Peces C, Pahl M O. Securing the Future of Railway Systems: A Comprehensive Cybersecurity Strategy for Critical On-Board and Track-Side Infrastructure[J]. Sensors, 2024, 24(24): 8218.

[12] Yu J, Wang R, Wu J. QoS-driven resource optimization for intelligent fog radio access network: A dynamic power allocation perspective[J]. IEEE

Transactions on Cognitive Communications and Networking, 2021, 8(1): 394-407.
[13] Wang X, Zhang Z, Hao Q, et al. Hardware-assisted security monitoring unit for real-time ensuring secure instruction execution and data processing in embedded systems[J]. Micromachines, 2021, 12(12): 1450.
[14] Ehret A, Del Rosario E, Gettings K, et al. A hardware root-of-trust design for low-power soc edge devices[C]//2020 IEEE High Performance Extreme Computing Conference (HPEC). IEEE, 2020: 1-6.
[15] Zhou Z, Liao H, Zhao X, et al. Reliable task offloading for vehicular fog computing under information asymmetry and information uncertainty[J]. IEEE Transactions on Vehicular Technology, 2019, 68(9): 8322-8335.
[16] Rekik M, Gransart C, Berbineau M. Analysis of security threats and vulnerabilities for train control and monitoring systems[C]//2018 15th International Multi-Conference on Systems, Signals & Devices (SSD). IEEE, 2018: 693-698.
[17] Verstichel S, Van Hoecke S, Strobbe M, et al. Ontology-driven middleware for next-generation train backbones[J]. Science of Computer Programming, 2007, 66(1): 4-24.
[18] Narouwa M, Mendiboure L, Badis H, et al. Enabling network technologies for flexible railway connectivity[J]. IEEE Access, 2024, 12: 151532-151553.
[19] Gala G, Fohler G, Tummeltshammer P, et al. RT-cloud: Virtualization technologies and cloud computing for railway use-case[C]//2021 IEEE 24th International Symposium on Real-Time Distributed Computing (ISORC). IEEE, 2021: 105-113.
[20] Feng Y, Zhong Z, Sun X, et al. Blockchain enabled zero trust based authentication scheme for railway communication networks[J]. Journal of Cloud Computing, 2023, 12(1): 62.
[21] Filipe A M A. Analysis of Security in Railway Communication Networks based on 5G and WiFi[J]. 2024.
[22] Saleh A M. Exploitation of 5G communications and positioning in future railway traffic management[J]. 2023.
[23] Lakshminarayana S, Karachiwala J S, Chang S Y, et al. Signal jamming attacks against communication-based train control: Attack impact and countermeasure[C]//Proceedings of the 11th ACM Conference on Security & Privacy in Wireless and Mobile Networks. 2018: 160-171.
[24] Liang H, Zhu L, Yu F R, et al. A cross-layer defense method for blockchain empowered CBTC systems against data tampering attacks[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 24(1): 501-515.
[25] Dabbaghzadeh H, Falahati A, Sanandaji N. CBTC security and reliability enhancements by a key-based direct sequence spread spectrum technique[J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 25(1): 159-172.
[26] Wang X, Liu L, Zhu L, et al. Joint security and QoS provisioning in train-centric CBTC systems under sybil attacks[J]. IEEE Access, 2019, 7: 91169-91182.
[27] Ren Q, Lin S, Cai Y, et al. Resource allocation and slicing strategy for multiple services co-existence in wireless train communication network[J]. IEEE Transactions on Wireless Communications, 2025, 24(1):401-414.
[28] Fadlullah Z M, Mao B, Kato N. Balancing QoS and security in the edge: Existing practices, challenges, and 6G opportunities with machine learning[J]. IEEE Communications Surveys & Tutorials, 2022, 24(4): 2419-2448.
[29] Zhang X, Niu Y, Yang T, et al. QoS-aware user association and transmission scheduling for millimeter-wave train-ground communications[J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(9): 9532-9545.
[30] Saquib N, Krintz C, Wolski R. Replicated versioned data structures for wide-area distributed systems[J]. IEEE Transactions on Parallel and Distributed Systems, 2022, 34(1): 207-224.
[31] Azad A, Selvitopi O, Hussain M T, et al. Combinatorial BLAS 2.0: Scaling combinatorial algorithms on distributed-memory systems[J]. IEEE Transactions on Parallel and Distributed Systems, 2021, 33(4): 989-1001.
[32] Zhang Q, Zhang Z, Cui J, et al. Efficient blockchain-based data integrity auditing for multi-copy in decentralized storage[J]. IEEE Transactions on Parallel and Distributed Systems, 2023, 34(12): 3162-3173.
[33] Dai H, Wang Y, Kent K B, et al. The state of the art of metadata managements in large-scale distributed file systems—scalability, performance and availability[J]. IEEE Transactions on Parallel and Distributed Systems, 2022, 33(12): 3850-3869.
[34] Gao Y, Gao X, Zhang R, et al. An end-to-end learning-based metadata management approach for distributed file systems[J]. IEEE Transactions on Computers, 2021, 71(5): 1021-1034.
[35] Zhang B, Kosar T. SMURF: Efficient and scalable metadata access for distributed applications[J]. IEEE Transactions on Parallel and Distributed Systems, 2022, 33(12): 3915-3928.

[36] Rapp M, Khalili R, Pfeiffer K, et al. Distreal: Distributed resource-aware learning in heterogeneous systems[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2022, 36(7): 8062-8071.
[37] Wang T, Du S, Cai H. CERT-DF: A computing-efficient and robust distributed deep forest framework with low communication cost[J]. IEEE Transactions on Parallel and Distributed Systems, 2023, 34(12): 3280-3293.
[38] Liu J, Zhang W, Ma T, et al. Toward security monitoring of industrial cyber-physical systems via hierarchically distributed intrusion detection[J]. Expert Systems with Applications, 2020, 158: 113578.
[39] Wu H, Peng Z, Guo S, et al. VQL: Efficient and verifiable cloud query services for blockchain systems[J]. IEEE Transactions on Parallel and Distributed Systems, 2021, 33(6): 1393-1406.
[40] Shalini D, Ashish K, Dhar A D, et al. Securing IoT devices: A novel approach using blockchain and quantum cryptography[J]. Internet of Things, 2024, 25: 101019.
[41] Zhang Q, Li Y, Wang R, et al. Data security sharing model based on privacy protection for blockchain-enabled industrial Internet of Things[J]. International Journal of Intelligent Systems, 2020, 36(1): 94-111.
[42] Khan A A, Laghari A A, Baqasah A M, et al. BDLT-IoMT—a novel architecture: SVM machine learning for robust and secure data processing in Internet of Medical Things with blockchain cybersecurity[J]. The Journal of Supercomputing, 2025, 81(1): 271.
[43] Guo T, Zhang Z, Yuan Y, et al. Hybrid concurrency control protocol for data sharing among heterogeneous blockchains[J]. Frontiers of Computer Science, 2024, 18(3): 183104.
[44] Yuan L, He Q, Tan S, et al. CoopEdge+: Enabling decentralized, secure and cooperative multi-access edge computing based on blockchain[J]. IEEE Transactions on Parallel and Distributed Systems, 2022, 34(3): 894-908.
[45] Wang Y, Li J, Zhao S, et al. Hybridchain: A novel architecture for confidentiality-preserving and performant permissioned blockchain using trusted execution environment[J]. IEEE access, 2020, 8: 190652-190662.
[46] Liu C, Guo H, Xu M, et al. Extending on-chain trust to off-chain–trustworthy blockchain data collection using trusted execution environment (tee)[J]. IEEE Transactions on Computers, 2022, 71(12): 3268-3280.
[47] Cabrera-Gutiérrez A J, Castillo E, Escobar-Molero A, et al. Integration of hardware security modules and permissioned blockchain in industrial iot networks[J]. IEEE Access, 2022, 10: 114331-114345.
[48] Guo Y, Zhang C, Wang C, et al. Towards public verifiable and forward-privacy encrypted search by using blockchain[J]. IEEE Transactions on Dependable and Secure Computing, 2022, 20(3): 2111-2126.
[49] Linoy S, Mahdikhani H, Ray S, et al. Scalable privacy-preserving query processing over ethereum blockchain[C]//2019 IEEE International Conference on Blockchain, 2019: 398-404.
[50] Khan A A, Laghari A A, Alroobaea R, et al. A lightweight scalable hybrid authentication framework for Internet of Medical Things (IoMT) using blockchain hyperledger consortium network with edge computing[J]. Scientific Reports, 2025, 15(1): 19856.
[51] Guo Y, Liu F, Zhou T, et al. Privacy vs. efficiency: achieving both through adaptive hierarchical federated learning[J]. IEEE Transactions on Parallel and Distributed Systems, 2023, 34(4): 1331-1342.
[52] Huang G, Wu Q, Sun P, et al. Collaboration in federated learning with differential privacy: A stackelberg game analysis[J]. IEEE Transactions on Parallel and Distributed Systems, 2024, 35(3): 455-469.
[53] Wang X, Wang J, Ma X, et al. A differential privacy strategy based on local features of non-gaussian noise in federated learning[J]. Sensors, 2022, 22(7):2424-2424.
[54] Li Y, Yang S, Ren X, et al. Multi-stage asynchronous federated learning with adaptive differential privacy[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 46(2): 1243-1256.
[55] Xu Y, Mao Y, Li S, et al. Privacy-preserving federal learning chain for internet of things[J]. IEEE Internet of Things Journal, 2023, 10(20): 18364-18374.
[56] Jing M, SiAhmed N, Stephan S, et al. Privacy-preserving federated learning based on multi-key homomorphic encryption[J]. International Journal of Intelligent Systems, 2022, 37(9):5880-5901.
[57] He Z, Wang L, Cai Z. Clustered federated learning with adaptive local differential privacy on heterogeneous iot data[J]. IEEE Internet of Things Journal, 2023, 11(1): 137-146.
[58] Ren C, Yu H, Yan R, et al. SecFedSA: A secure differential privacy-based federated learning approach for smart cyber-physical grid stability assessment[J]. IEEE Internet of Things Journal,

2024, 11(4): 5578-5588.
[59] Wei C, Yu R, Fan Y, et al. Securely sampling discrete gaussian noise for multi-party differential privacy[C]//Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security. 2023: 2262-2276.
[60] Zhang Y, Li S, Yang L. Distributed optimal control to virtual formation of railway trains with dynamic coupling/decoupling: An accelerated projected gradient based decomposition method[J]. IEEE Transactions on Vehicular Technology, 2023, 72(12): 15405-15420.
[61] Zhang Z, Li J, Sun Y, et al. Blockchain-based secure key management model for high-speed railway[C]//2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2022: 1988-1993.
[62] Liang H, Zhu L, Yu F R. Collaborative edge intelligence service provision in blockchain empowered urban rail transit systems[J]. IEEE Internet of Things Journal, 2024, 11(2): 2211-2223.
[63] Qin N, Du J, Zhang Y, et al. Fault diagnosis of multi-railway high-speed train bogies by improved federated learning[J]. IEEE Transactions on Vehicular Technology, 2023, 72(6): 7184-7194.
[64] Jiang S, Cao J, Wu H, et al. Privacy-preserving and efficient data sharing for blockchain-based intelligent transportation systems[J]. Information Sciences, 2023, 635: 72-85.
[65] Lim H W, Temple W G, Tran B A N, et al. Data integrity threats and countermeasures in railway spot transmission systems[J]. ACM Transactions on Cyber-Physical Systems, 2019, 4(1): 1-26.
[66] Vahidi V. Uplink data transmission for high speed trains in severe doubly selective channels of 6G communication systems[J]. Physical Communication, 2021, 49: 101489.
[67] Wang Z, Xie X, Chen L, et al. Intrusion detection and network information security based on deep learning algorithm in urban rail transit management system[J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(2): 2135-2143.
[68] Du J, Cheng J, You Y, et al. Federated transfer learning for fault diagnosis of high-speed train bogie with data security and training optimization[C]//2023 CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes, {SAFEPROCESS}, 2023:1-6.
[69] Saki M, Abolhasan M, Lipman J, et al. A comprehensive access point placement for iot data transmission through train-wayside communications in multi-environment based rail networks[J]. IEEE Transactions on Vehicular Technology, 2020, 69(10): 11937-11949.
[70] Wang K, Li H, Zhang Q. Parallel redundancy protocol for railway wireless data communication network[J]. Wireless Communications and Mobile Computing, 2022.
[71] Chan R. A security framework for railway system deployments[C]//Critical Infrastructure Protection XV: 15th IFIP WG 11.10 International Conference, 2022: 247–253.
[72] Luo M, Zhu L. Research on real-time and reliability of wireless transmission of high-speed train control data based on data mining technology[C]//Journal of Physics: Conference Series. IOP Publishing, 2021, 70(4): 042093.
[73] Soderi S, Masti D, Lun Y Z. Railway cyber-security in the era of interconnected systems: a survey[J]. IEEE Transactions on Intelligent Transportation Systems, 2023. 24(7): 6764-6779.
[74] Wang X, Li C, Wu W, et al. Energy balanced data transmission strategy for LWSN in railway environment monitoring[C]//Proceedings of the 2020 8th International Conference on Information Technology: IoT and Smart City. 2020: 189-194.
[75] Wu L. Design of data transmission system for 3D laser scanning of liquefied gas railway tanker based on fuzzy algorithm[J]. Journal of Intelligent & Fuzzy Systems, 2020, 38(6): 7755-7766.