Dean&Francis
ISSN 2959-6157

A Survey of Research Advances in Federated Learning for Anomaly Detection

Futao Ye

School of Artificial Intelligence, South China Normal University, Guangzhou, China
*Corresponding author: yefutao@m.scnu.edu.cn

Abstract:

In fields such as industrial control systems, cybersecurity, and finance, detecting abnormal behaviors is of utmost importance to ensure operational reliability and prevent potential losses. In recent years, the advantage of Federated Learning (FL) in addressing privacy protection challenges and enabling collaborative model training across distributed data sources has made it an increasingly important solution for anomaly detection tasks. This paper provides a comprehensive review of the application of FL in various anomaly detection scenarios, including cybersecurity, intelligent industry, and financial services. It systematically analyzes the performance of popular models such as RNN/LSTM, AE/VAE, GNN, and Transformer-based methods under non-IID data distributions common in federated settings. Furthermore, this study identifies and discusses several critical challenges faced in FL-based anomaly detection, such as data heterogeneity, communication overhead, robustness, and security threats. Corresponding strategies to mitigate these issues are also summarized. Experimental evidence and case studies demonstrate that FL can achieve efficient, privacy-preserving, and scalable distributed anomaly detection, and thus holds significant promise for deployment in real-world applications.

Keywords: Federated Learning; Anomaly Detection; Unsupervised Learning; Model Compression.

1. Introduction

Centralized machine learning is a type of machine learning where all the data is centralized to a central location for training. However, privacy issues and strict regulatory requirements are difficult challenges for it to overcome. This is especially true in areas such as finance, healthcare and the industrial Internet of Things, where data is often controlled by various organizations. This causes cooperation to become difficult and slows down the process of building effective models. It can be seen that privacy protection has become a key factor which is affecting the application of artificial intelligence technology.

FL has been widely used in various fields [1,2]. It enables multiple clients to collaboratively train a global model while maintaining data localization and user privacy integrity. The main idea is to conduct local training on different devices, then upload the updates

ISSN 2959-6157

and aggregate and update the model on the central server to protect data privacy. It is worth noting that, unlike traditional methods, FL achieves a natural balance between model accuracy and privacy protection, making it highly suitable for collaborative tasks across different domains, devices, and organizations. With the development of The Times, the infrastructure such as edge computing, Internet of things devices, and 5G networks have been continuously improved, and federated learning has found applications in areas such as smart devices, wearables, autonomous vehicles, financial risk management, and healthcare [3], becoming a major topic in both research and industry. In addition, thanks to the continuous improvement of federated optimization algorithms such as FedAvg, FedProx, FedNova, and scaffolding, FL is also robust and effective under challenging network conditions and heterogeneous hardware settings.

As an important means of system identification of abnormal behavior, anomaly detection has a wide range of applications in network security, industry, finance and other fields. Traditional methods rely on centralized data collection, which also leads to difficulties in traditional methods on issues such as privacy, data heterogeneity and limited labels. Moreover, in large-scale complex networks, centralized methods are prone to problems such as delay, bottleneck and vulnerability to attack. In addition to this, when data is Non-Independent identically distributed(Non-IID) or changes dynamically, centralized models often perform poorly, causing false alarms or missed detections.

Federated learning, owing to its capability of handling heterogeneous data and unique advantages, is regarded as a preferred solution in domains such as intelligent manufacturing, smart homes and finance. The nature of federated detection also allows for distributed threat awareness and defense.Such a property greatly reduces the dependence on a centralized system. With deeper research on other aspects, such as self-supervised learning, neural networks and model compression, it is also continuously improving accuracy, efficiency and security, showing strong practical potential.

Although federated learning for anomaly detection is still evolving rapidly, it also faces many technical difficulties. Challenges include handling unstructured data [4], improving communication efficiency, managing complex system deployments and defending against malicious participants. Additionally, it is noticeable that factors in the real world such as data synchronization delays and fluctuations in device participation also increase the difficulty of deployment of FL and raise higher requirements for the robustness of the system. From the above, how to improve the stability, scalability and privacy protection of anomaly detection while maintaining strong performance is a focus to be solved.

Therefore, aiming to help researchers gain a comprehensive understanding of the current situation and future development direction of this field, this paper systematically reviews the latest research progress of federated learning in the field of anomaly detection.

2. Research Status of Federated Learning in Anomaly Detection

2.1 Typical Application Scenarios

2.1.1 Network Security

To deal with the Non-IID challenge prevalent in IoT scenarios, Li et al. designed a clustered federated learning architecture [5]. By calculating the similarity of data distribution and characteristics between devices, the architecture dynamically forms multiple clusters and carries out personalized training within the cluster, so that the global model can better adapt to different types of devices. Experimental results show that compared with the traditional FedAvg method, the proposed grouping strategy significantly alleviates the model performance problems caused by data heterogeneity. It not only improves anomaly detection accuracy but also enhances the robustness and stability of the system. Experimental results demonstrate that this architecture exhibits clear advantages in various real-world IoT scenarios, particularly excelling at identifying complex and stealthy network attacks, indicating strong potential for deployment in large-scale heterogeneous device environments.

Alsulaimawi, from a security perspective, integrated autoencoder reconstruction errors with gradient analysis in federated learning to design a hybrid anomaly detection framework that significantly improves the robustness of local models and enhances resistance against model poisoning attacks [6]. By leveraging both the reconstruction errors-indicating data anomalies-and gradient information-revealing potential malicious updates-this method constructs a multi-dimensional anomaly scoring mechanism to accurately identify and isolate malicious clients. Experiments conducted on multiple public datasets, including image and time-series data, show that the framework not only boosts the security performance of federated anomaly detection systems but also effectively mitigates various attacks targeting federated learning without compromising model performance. This marks a significant step in securing FL-based anomaly detection.

2.1.2 Industrial Systems

Husnoo et al. introduced a decentralized P2P gossip protocol-driven federated learning mechanism tailored for

anomaly detection in smart grids [7]. This approach replaces the traditional centralized aggregator with peer-to-peer communication, effectively eliminating single points of failure and lowering privacy concerns. The design has achieved remarkable results in enhancing system stability and privacy protection. The team proposed Decentralized Federated Learning (DFL) + Transformer-AutoEncoder (TAE) and introduced DP-SIGNSGD gradient quantization, which reduced the training time by around 35% and achieved high accuracy. It can be seen that the proposed architecture can flexibly adapt to the diverse characteristics of the distributed industrial environment and meet the stringent requirements for real-time response and communication overhead. In the smart grid related experiments, the results show that the proposed model effectively reduces the communication cost while maintaining a high detection accuracy, and significantly improves the overall operation efficiency and security of network anomaly detection.

2.1.3 Financial Sector

To address the issue of extremely sensitive and non-centrally-storable data in the IoT heterogeneous environment for financial fraud detection, this team specially designed an aggregation method based on category grouping - FedGroup [8]. This method divides nodes into homogeneous groups according to device characteristics, and then performs data aggregation within the groups, thereby solving the Non-IID problem in federated learning. In addition, the FedGroup framework integrates ensemble learning techniques, which can resist certain noise and attacks, reduce training fluctuations, and improve the model's generalization ability. Comprehensive experimental evaluations show that the proposed solution performs well in complex and dynamic financial fraud detection tasks. Considering high accuracy and strict privacy protection, this method provides reliable technical support for the application of federated anomaly detection in large-scale, heterogeneous, and sensitive data environments.

2.2 Methodologies and Technical Approaches

The federated learning methods used for anomaly detection can be broadly classified into the following categories: autoencoder techniques including AE, temporal models based on recurrent neural networks (RNNs) or long short-term memory networks (LSTMs) and VAE. Each method has its unique advantages and is capable of addressing the challenges posed by privacy requirements, different data formats, and limited computing resources.

2.2.1 RNN/LSTM-based Models

Due to their excellent performance in handling sequential data and capturing time dependencies, RNNs and LSTMs have been widely applied in the federated anomaly detection framework. Zhang et al. proposed a federated time series anomaly detection framework that combines the variational autoencoder (VAE) with the long short-term memory network to fully utilize the advantages of these two models. This architecture achieves precise detection of complex and subtle anomalies in Non-IID environments by integrating temporal models on edge devices with distributed learning, local adaptive training, robust aggregation and trust evaluation mechanisms under federated learning. In addition, this model architecture demonstrates strong generalization ability, which has shown certain practicality and reliability in complex application scenarios such as intelligent industry and Internet of Things [9]. In this approach, lightweight LSTM models are deployed on local devices to ensure computational efficiency, while privacy-preserving parameter updates are aggregated on the central server, thereby jointly constructing a powerful global model. This design effectively addresses the challenges commonly encountered in real-world federated learning scenarios involving Non-IID multivariate time series data.

Furthermore, to mitigate training instability and gradient vanishing issues commonly associated with LSTM in federated settings, some studies have introduced residual connections, attention mechanisms, or sliding window techniques to enhance detection accuracy and robustness.

2.2.2 AE/VAE-based Models

Autoencoders (AE) and their variants have gradually become mainstream in federated anomaly detection. Xu et al. proposed a federated VAE framework that integrates parameter-efficient fine-tuning (PeFT) and knowledge distillation. To address resource limitations on edge devices, the approach freezes most network parameters and performs fine-tuning only on a few "adapter layers," thereby significantly reducing communication and computational overhead, while improving performance using a teacher-student model structure [10]. Experimental results show broad applicability in scenarios such as smart factories and wearable devices.

Another representative work by Nardi et al. adopted an unsupervised AE framework with client-side clustering. By conducting localized updates within clusters rather than sharing parameters directly, this approach better accommodates the heterogeneous nature of multi-source data [11]. It is particularly well-suited for unlabeled scenarios such as network intrusion detection and financial transaction monitoring.

In addition, some studies have enhanced the representation power and improved the discriminability of anomalies by combining variational autoencoders with contrastive learning. They also adopted a federated optimizer

ISSN 2959-6157

with dynamic weighting capabilities. By implementing this function during the local model aggregation process, they enhanced the overall stability.

2.2.3 Other Approaches

In the latest advancements of federated anomaly detection, Zhang et al. have demonstrated particularly outstanding performance [12]. They introduced a federated detection framework based on GNN specifically designed for vehicle CAN bus systems. They modeled the CAN bus as a graph, enabling the system to learn the complex spatial dependencies and temporal dynamics in the communication signals. This design not only improves the detection accuracy of complex attack patterns but also meets the requirements of data locality and privacy protection. By focusing on this graph-based federated framework, Zhang et al. provided a powerful solution that addresses the unique challenges of vehicle networks and distinguishes it from other mainstream methods.

3. Core Challenges and Research Trends

3.1 Core Challenges

Dealing with Non-IID data is always a challenging task in federated learning. In many practical scenarios, due to differences in device, client data often varies significantly. This diversity hinders the convergence of the global model and reduces its generalization ability. To address this issue, various strategies have been introduced, including training multiple models, clustering clients based on their model updates, and optimizing models to achieve personalization. For example, unsupervised device clustering techniques performed within the federated learning framework have improved anomaly detection performance in heterogeneous IoT networks by grouping devices with similar behavior patterns [5]. Despite the progress of these solutions, they can still be vulnerable to the quality of local data and may face difficulties in adversarial contexts. Therefore, developing communication-efficient, adaptive and robust methods to address Non-IID data remains a key focus of future research.

Large-scale deployment of federated learning still needs to consider the limitations of computing resources and communication efficiency. The frequent information interaction caused by model updates will bring huge communication overhead, which is particularly prominent in edge devices and IoT environments. To address these challenges, various techniques have emerged as research hotspots, including weight quantization, model compression, parameter-efficient fine-tuning (PeFT), and sparsification. These approaches strive to drastically reduce communication payloads and computational costs without sacrificing model accuracy, leading to more scalable and responsive federated learning systems that are more applicable to real-world heterogeneous and resource-constrained scenarios [6].

Privacy protection and system robustness challenges remain critical. Security risks such as man-in-the-middle attack and malicious node model poisoning may still seriously affect the reliability of the federated learning model. In order to enhance the system defense capability, designing effective mechanisms to detect and isolate malicious participants is still a key problem to be studied.

The related devices involved in FL often vary greatly in terms of hardware specifications, data formats, and task types, which complicates stable training and limits the generalization ability of the global model. Clustered federated learning approaches that group similar devices based on their communication behaviors or model characteristics have shown better convergence and personalized results in large heterogeneous IoT networks [5]. Additionally, parameter-efficient federated learning frameworks tailored for multi-task settings help overcome computational limits and support a range of applications [9]. Consequently, creating flexible and scalable federated learning systems capable of effectively managing device heterogeneity and multi-task collaboration remains a crucial area for future research.

3.2 Research Trends and Future Directions

Unsupervised and lightly supervised federated anomaly detection is gradually becoming a main research focus. Traditional anomaly detection methods usually depend on large amounts of labeled data, but in practice, anomalous samples are rare and labeling is costly. Therefore, unsupervised methods (such as AE/VAE based on reconstruction errors and adaptive clustering) and lightly supervised strategies (like pseudo-label generation and self-supervised contrastive learning) are especially important within federated frameworks. Future research will focus on constructing more robust global models to handle data heterogeneity, designing federated self-supervised tasks to exploit local client structures, and developing federated evaluation metrics suited for unlabeled scenarios. Moreover, multimodal fusion (e.g., images combined with sensor data) and cross-domain knowledge transfer will help improve model generalization and practical adaptability.

Model compression and acceleration technologies are key to promoting practical federated learning deployment. As resource constraints on terminal devices (computation, storage, bandwidth) become more severe, the demand for lightweight federated models is growing. Current mainstream approaches include local model pruning, weight quantization for transmission compression, distillation-as-

sisted model transfer and sharing, and parameter-efficient fine-tuning (PeFT) via low-rank representations such as LoRA. Future studies will further focus on how to dynamically adjust compression rates to balance accuracy and resource consumption, design joint communication-computation optimization strategies, and integrate with differential privacy mechanisms to enhance overall system flexibility and controllability.

Integration of Graph Neural Networks (GNNs) and Transformers provides new ideas for modeling complex structured data. In practical anomaly detection scenarios (such as intelligent transportation and industrial IoT), data often exhibits strong spatiotemporal coupling and node dependencies. GNNs effectively capture high-order relationships among topologies, while Transformers excel in sequence modeling. Their fusion (e.g., Graphormer, GTN) enables models to learn both local dependencies and global attention mechanisms, significantly improving detection accuracy and generalization. Designing efficient algorithms that support local graph structure updates, cross-client graph sharing, and heterogeneous attention fusion under federated settings will become a research hotspot.

Introduction of blockchain and trustworthy collaboration mechanisms is building new security architectures for federated learning. In multi-party federated systems, the authenticity of model updates, client trustworthiness, and overall fairness of collaboration are major challenges. Blockchain technology, with its immutability, distributed ledger, and smart contract capabilities, can realize transparent auditing of the training process and automated punishment rules. Future directions include designing on-chain model aggregation mechanisms, combining with differential privacy and secure multi-party computation (SMC) to achieve multi-dimensional defense, and optimizing low-latency, high-throughput blockchain architectures for federated systems.

4. Conclusion

This paper reviews recent advances in federated learning for anomaly detection, highlighting key applications including network security, industrial monitoring, and financial fraud detection. It examines major models and technical approaches, demonstrating how federated learning effectively overcomes data silos and privacy concerns, thereby enhancing both the security and practicality of anomaly detection systems.

Despite these advances, challenges such as Non-IID data modeling, communication overhead, vulnerability to malicious attacks, and limited generalization in heterogeneous environments persist. Future research directions include focusing on unsupervised learning methods, model compression and lightweight techniques, multimodal data fusion, and cross-domain knowledge transfer. Furthermore, integrating emerging technologies like graph neural networks, Transformers, and blockchain holds promise for advancing the field.

In summary, federated learning presents an innovative, privacy-preserving, and distributed collaborative paradigm with broad application potential in anomaly detection. This review aims to provide guidance for ongoing research and facilitate practical implementations.

References

[1] Aledhari M, Razzak R, Parizi R M, et al. Federated learning: a survey on enabling technologies, protocols, and applications. IEEE Access, 2020, 8: 140699–140725.

[2] Li L, Fan Y, Tse M, Lin K Y, et al. A review of applications in federated learning. Computers and Industrial Engineering, 2020, 149: 106854.

[3] Rieke N, Hancox J, Li W, Milletari F, et al. The future of digital health with federated learning. NPJ Digital Medicine, 2020, 3(1): 119.

[4] Lim W Y B, Luong N C, Hoang D T, et al. Federated learning in mobile edge networks: a comprehensive survey. IEEE Communications Surveys & Tutorials, 2020, 22(3): 2031–2063.

[5] LI Xiaoming, WANG Yifan, ZHANG Tao. Clustered federated learning architecture for network anomaly detection in large-scale heterogeneous IoT networks. Computers & Security, Elsevier, 2023, 126: 103299.

[6] Alsulaimawi Z. Anomaly detection via gradient and autoencoder analysis in federated learning. arXiv preprint arXiv:2403.10000, 2024.

[7] Husnoo M A, Anwar A, Haque M E, et al. Decentralized federated anomaly detection in smart grids: A P2P gossip approach. arXiv:2407.15879v2, 2025-01-09.

[8] Zhang Y, Suleiman B, Alibasa M J, Farid F. Privacy-aware anomaly detection in IoT environments using FedGroup: a group-based federated learning approach. Journal of Network and Systems Management, 2024, 32(1): 20.

[9] Zhang C, Yang S, Mao L, Ning H. Anomaly detection and defense techniques in federated learning: a comprehensive review. Artificial Intelligence Review, 2024, 57(6): 150.

[10] Xu R, Miao H, Wang S, Yu P S, Wang J. PeFAD: A parameter-efficient federated framework for time series anomaly detection. arXiv preprint arXiv:2406.02318, 2024.

[11] NARDI Mirko, VALERIO Lorenzo, PASSARELLA Andrea. Anomaly detection through unsupervised federated learning. arXiv preprint, arXiv:2209.04184, 2022-09-09.

[12] ZHANG Hengrun, ZENG Kai, LIN Shuai. Federated Graph Neural Network for Fast Anomaly Detection in Controller Area Networks. IEEE Transactions on Information Forensics and Security, IEEE, 2023, 18(5): 1566-1579.