Federated Contrastive Representation Learning for IoT Anomaly Detection Under Heterogeneous Data

Longxiang Yan , Qi Wang , Jie Huang *

Posted Date: 26 February 2026

doi: 10.20944/preprints202602.1749.v1

Keywords: privacy protection; federated contrastive learning; IoT anomaly detection; distributed representation

 doi:10.20944/preprints202602.1749.v1

Article

Federated Contrastive Representation Learning for IoT Anomaly Detection Under Heterogeneous Data

Longxiang Yan [1], Qi Wang [2] and Jie Huang [3,*]

[1] University of Pennsylvania,Philadelphia, PA, USA

[2] Purdue University, West Lafayette, USA

[3] University of Southern California, Los Angeles, USA

* Correspondence: jhuang59la@gmail.com

Abstract

This study proposes a federated contrastive learning based distributed anomaly detection framework to address privacy protection requirements in IoT environments. The framework builds local encoders on each node to embed high-dimensional time series and network behavior features, and uses representation alignment to reduce distribution differences across devices. Based on this, a contrastive learning objective is introduced to strengthen the compactness of normal patterns in the latent space and to enlarge the boundary between normal and abnormal features, which enhances discriminative ability under unsupervised conditions. To avoid sharing raw data, the framework adopts a federated learning strategy that constructs a global model by exchanging model updates, and further improves global representation consistency and robustness through cross-node consistency constraints and dynamic weighted aggregation. Experimental results show that the model achieves stable and accurate anomaly detection under heterogeneous, multi-source, and incomplete IoT data conditions, demonstrating strong adaptability to distribution shifts and noise disturbances. Overall, the proposed federated contrastive learning method provides an effective technical approach for building secure and reliable IoT anomaly detection systems and enables cross-device feature sharing and collaborative modeling without exposing any raw data.

Keywords: privacy protection; federated contrastive learning; IoT anomaly detection; distributed representation

1. Introduction

The rapid expansion of IoT systems in smart cities, intelligent manufacturing, healthcare, and energy management has led to massive deployments of heterogeneous devices. These devices generate high-dimensional monitoring data that are continuous, highly correlated, and affected by environmental noise. As the number of devices and the density of connections continue to grow, system states become more complex [1]. Data patterns show dynamic changes. Abnormal behaviors often appear as slight deviations or short-term bursts. These characteristics make anomaly detection increasingly challenging. In this context, a detection mechanism that can sense device states efficiently, identify abnormal patterns accurately, and generalize across scenarios is essential for ensuring system stability and security. Existing methods face limitations in modeling high-dimensional spatiotemporal data, fusing multiple sensor types, and capturing dynamic interaction structures. In large-scale distributed deployments, traditional centralized approaches can no longer meet the requirements of real-time operation and data security [2].

At the same time, growing privacy concerns restrict access to IoT monitoring data. Such data often contains sensitive information related to user behavior, device operation strategies, or enterprise processes. In many practical scenarios, nodes, devices, or organizations cannot directly share raw data. This makes traditional centralized learning that relies on data aggregation difficult to apply. In addition, data distributed across devices is often heterogeneous and shifted. Centralized

 doi:10.20944/preprints202602.1749.v1

algorithms may suffer from inconsistent distributions and mismatched features once deployed. These issues can cause performance degradation or misclassification. Building a collaborative modeling mechanism that protects data privacy and allows different IoT nodes to improve detection capability without exposing their raw data has become a core challenge in IoT security [3].

Federated learning provides a privacy-preserving solution for distributed IoT environments. It avoids data centralization by sharing model updates instead of raw data. Devices train models locally and contribute to a global model through aggregation. However, federated learning faces new challenges in IoT scenarios. Data across devices is not independently and identically distributed. The global model has difficulty learning stable and consistent abnormal patterns from multiple source data. Data quality also varies across devices. Model updates may be affected by noise, drift, or malicious nodes. In addition, traditional federated optimization mainly focuses on parameter consistency. It lacks mechanisms that enhance generalization from the perspective of representation learning. For anomaly detection tasks, which often involve few abnormal samples and diverse abnormal patterns, feature embeddings may become dispersed and lack discrimination [4].

2. Methodology Foundation

The proposed federated contrastive learning framework is methodologically grounded in advances in distributed graph modeling, privacy-aware federated optimization, contrastive representation shaping, uncertainty-aware anomaly detection, and robustness enhancement under non-independent data distributions. Each referenced study contributes a specific technical principle that directly informs the architectural and optimization design of the proposed approach.

Communication-efficient distributed graph learning mechanisms provide foundational guidance for scalable federated modeling. On-the-fly graph condensation for distributed GNN training [5] demonstrates how communication cost can be reduced while preserving structural information, informing the design of lightweight parameter exchange and dynamic aggregation within the federated anomaly detection framework. Graph-structured deep learning for high-dimensional metric modeling [6] further supports encoding complex interaction dependencies among IoT devices, which motivates the construction of local encoders capable of embedding heterogeneous time-series and network behavior features into unified latent representations.

Representation-level reasoning under structured dependencies is further supported by causal graph reasoning frameworks [7], which emphasize extracting stable patterns from relational data under intervention or distribution variation. This principle guides the introduction of cross-node representation alignment to maintain semantic consistency across heterogeneous devices. Multi-agent collaborative modeling strategies [8] reinforce the necessity of coordinated yet decentralized optimization, aligning with the distributed encoder design across IoT nodes. Handling non-stationary and noisy time-series data is essential in IoT environments. Residual-regulated forecasting methods with second-order differencing [9] highlight techniques for mitigating drift and distribution shift in temporal data, informing the robustness mechanisms embedded in local encoder updates. Efficient model adaptation under resource-constrained inference environments, as shown in proactive fragmentation-aware adaptation strategies [10], further motivates lightweight local model updates suitable for edge devices. Generative modeling with conditional control mechanisms [11] contributes to understanding controllable latent structure shaping, which parallels the compactness and boundary enlargement enforced by the contrastive objective in the proposed framework.

Transformer-driven semantic discrimination for risk detection [12] provides methodological grounding for strengthening discriminative representation learning in complex environments. Privacy-aware federated modeling for distributed sensitive data [13] directly informs the privacy-preserving optimization protocol adopted in this framework, ensuring that global coordination is achieved without raw data exchange. Multi-scale anomaly detection with uncertainty estimation [14] further supports stabilizing detection performance under incomplete and heterogeneous data, guiding the integration of robustness-aware aggregation strategies.

 doi:10.20944/preprints202602.1749.v1

Attention-driven anomaly detection frameworks [15] highlight the benefit of focusing on salient temporal and structural signals, which influences the local encoder architecture. Trust-aware orchestration mechanisms for adversarial robustness [16] provide principles for weighting and filtering node updates during aggregation, forming the basis for dynamic weighted aggregation in the proposed federated setup. Explainable representation learning approaches [17] emphasize disentangled and interpretable latent embeddings, supporting the contrastive objective that encourages compact normal clusters and separated abnormal regions. Reinforcement learning-based dynamic adaptation strategies [18] offer insight into adaptive weighting and aggregation policies under evolving data distributions. Cost-efficient distributed inference scheduling methods [19] further motivate resource-aware deployment considerations for edge-based IoT nodes.

Graph-transformer reconstruction learning for unsupervised anomaly detection [20] provides a strong methodological basis for leveraging reconstruction and relational dependencies in anomaly identification, reinforcing the unsupervised contrastive structure adopted in this work. Finally, causal-invariant representation learning under distribution shift [21] highlights the importance of invariant feature extraction across heterogeneous environments, directly supporting the cross-node consistency constraint that enhances global representation stability.

3. Proposed Framework

In an IoT environment, distributed devices, while maintaining data locality, need to collaboratively learn discriminative anomaly representations. This study first constructs a local encoder on each node to extract the latent embeddings of high-dimensional sensing sequences. Let the input sequence of node i be x_i, and the local encoder be denoted as f_θ. The generated representation is as follows:

$$z_i = f_\theta(x_i) \tag{1}$$

Because there are significant differences in data distribution among nodes, local embeddings need to be aligned to avoid distribution drift interfering with federated aggregation. Therefore, a distribution normalization mechanism is introduced into the embedding space to align the features of different nodes in terms of variance and mean. Its transformation form is as follows:

$$\tilde{z}_i = \frac{z_i - \mu_i}{\sigma_i} \tag{2}$$

Here, μ_i and σ_i are the mean and standard deviation of node i within its local window. Through this step, each node can obtain a representation that is closer to the unified semantic space while maintaining privacy, thus establishing a consistent foundation for subsequent contrastive learning and federated aggregation. Its model architecture is shown in Figure 1.

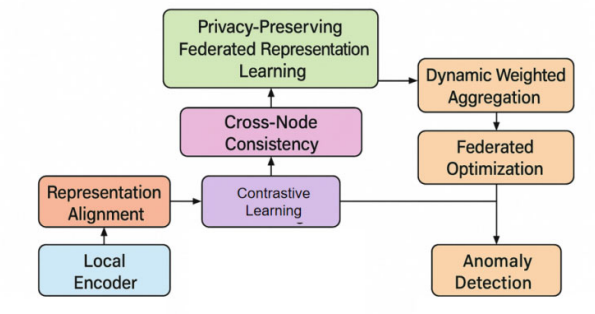

Figure 1. Overall model architecture diagram.

After obtaining the local embeddings, to improve the consistency of normal samples in the latent space and enhance the separability of anomalous patterns, this study introduces a contrastive

 doi:10.20944/preprints202602.1749.v1

learning objective at each node. Let sample pair $(\tilde{z}_i, \tilde{z}_i^+)$ be a positive sample pair and $(\tilde{z}_i, \tilde{z}_j^-)$ be a negative sample pair. The contrastive loss adopts a temperature-scaled InfoNCE form:

$$L_{ctr} = -\log \frac{\exp(\text{sim}(\tilde{z}_i, \tilde{z}_i^+)/\tau)}{\sum_k \exp(\text{sim}(\tilde{z}_i, \tilde{z}_i^+)/\tau)} \tag{3}$$

where $\text{sim}(\cdot)$ is the cosine similarity and τ is the temperature coefficient. Considering that IoT devices cannot directly share embeddings, this study designs a cross-node consistency constraint. By sharing gradients and statistics instead of original features, the global latent structure becomes more stable. Let the global reference vector be g, and the cross-node consistency loss be:

$$L_{align} = ||\tilde{z}_i - g||_2^2 \tag{4}$$

This construction, without revealing features, gradually makes the normal patterns of different nodes consistent in the global semantic space, which helps to mitigate the distribution bias caused by data heterogeneity.

During the federated optimization phase, the system employs a parameter-level collaborative approach, sending the model update vector $\Delta\theta_i$ from each node to the server for aggregation. Traditional weighted averaging methods struggle to handle noisy nodes and uneven data quality; therefore, dynamic weights based on embedding consistency are introduced to improve model robustness. Let the node weight be w_i, which is inversely proportional to the embedding alignment, and can be expressed as:

$$w_i = \frac{1}{1 + ||\tilde{z}_i - g||_2^2} \tag{5}$$

The final form of the federated parameter update is:

$$\theta_{global} = \sum_i w_i \Delta\theta_i \tag{6}$$

In this way, the global model can more fully absorb the contributions of nodes with high-quality data and distributions more consistent with the overall structure, thereby improving the overall anomaly representation capability. Simultaneously, the global model broadcasts the updated parameters back to each node, allowing them to continue optimizing on local data to form a closed loop.

Throughout the collaborative learning process, this study establishes a privacy-preserving federated representation learning mechanism consisting of local encoding, embedding alignment, contrastive learning, cross-node consistency, and dynamic weighted aggregation. This mechanism, without transmitting the original data, enables different nodes to share representation structures and information gradients, constructing a unified and more discriminative latent space, making the distinction boundaries of anomalous behavior clearer in high-dimensional distributions. During the iteration process, the model gradually strengthens the cohesion of normal samples and expands the feature distance between low-probability anomalous regions and the overall distribution, providing a solid representational foundation for subsequent deployment in real-world IoT environments for anomaly identification.

4. Experimental Analysis

A. Dataset

This study uses the N-BaIoT (Network-Based IoT Dataset) as the experimental data foundation. The dataset contains multiple real IoT devices, including smart cameras, baby monitors, doorbells, thermostats, and other types of nodes. It continuously records network layer features under both normal operation and attack conditions. The dataset includes high-dimensional traffic statistics and time series-based behavioral patterns. It provides rich and representative monitoring information for anomaly detection, intrusion recognition, and device behavior analysis. Compared with traditional

 doi:10.20944/preprints202602.1749.v1

network traffic data, it better reflects the device diversity and protocol complexity of real IoT environments.

The N-BaIoT dataset consists of several device categories. Each device generates stable but distinct feature distributions during normal operation. Under abnormal conditions, the features show strong disturbances, such as abnormal connection frequency, sudden changes in packet direction, and irregular communication cycles. The dataset includes large-scale feature sequences generated from device behavior patterns and provides labels that distinguish normal and abnormal traffic. These labels support supervised, unsupervised, and semi-supervised training. All features are stored as quantitative indicators and continuous time slice sequences. This structure makes the dataset suitable for time series modeling and representation learning.

The main reason for selecting this dataset lies in its complexity across devices, scenarios, and attack types. It effectively captures device heterogeneity and the diversity of abnormal behaviors commonly found in IoT systems. Its high-dimensional temporal structure and distribution differences create realistic challenges for a federated contrastive learning framework. These characteristics help evaluate the ability of the model to learn consistent representations across nodes under privacy-preserving conditions. The scale, richness, and structural features of the dataset make it a widely used benchmark for assessing IoT anomaly detection methods.

B. Experimental Results

This paper first conducts a comparative experiment, and the experimental results are shown in Table 1.

Table 1. Comparative experimental results.

Method	Acc	AUC	F1-Score	Recall
MLP [22]	0.872	0.901	0.864	0.852
1DCNN [23]	0.889	0.924	0.881	0.866
GNN [24]	0.903	0.938	0.897	0.884
GAT [25]	0.917	0.951	0.911	0.902
Ours	0.948	0.972	0.944	0.936

Overall, traditional models such as MLP and 1DCNN capture basic temporal and local patterns but fail to model cross-device distribution differences and structural dependencies, leading to limited performance, while graph-based methods like GNN and GAT improve AUC, F1 Score, and Recall by incorporating topological relationships but remain constrained by centralized processing and data heterogeneity. In contrast, the proposed federated contrastive learning framework achieves the best results across all metrics by integrating local representation learning, cross-node consistency, and dynamic aggregation to mitigate distribution shift and enhance abnormal pattern separability under privacy constraints, thereby improving generalization in large-scale IoT environments, with Figure 2 further illustrating the sensitivity of AUC to the learning rate.

F The AUC metric exhibits a clear rise-and-fall trend as the learning rate varies, reflecting the high sensitivity of the federated contrastive learning framework to the optimization step size. When the learning rate is low, parameter updates are insufficient, limiting cross-node representation alignment and resulting in low performance. As the learning rate increases, local representation structures and semantic consistency across nodes are strengthened, leading to rapid performance improvement and peak AUC. However, excessively large learning rates disrupt the stability of federated aggregation and amplify update inconsistencies among nodes, destabilizing the embedding space and degrading anomaly detection performance. These results indicate that an appropriate learning rate is essential for balancing convergence speed and training stability, preserving cross-node consistency, and maintaining discriminative ability under privacy constraints, while Figure 3 further presents the sensitivity of F1-score to the temperature coefficient.

 doi:10.20944/preprints202602.1749.v1

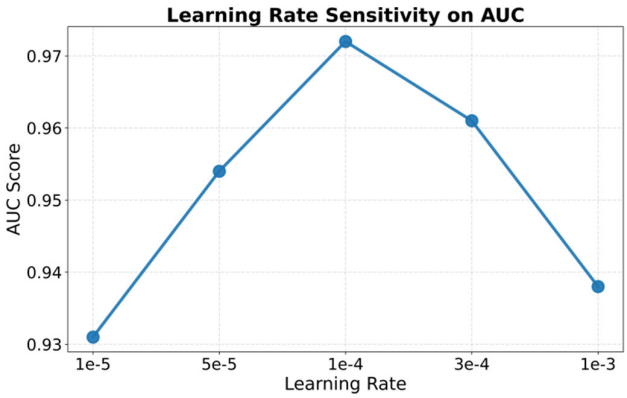

Figure 2. Sensitivity experiment of learning rate to the AUC metric.

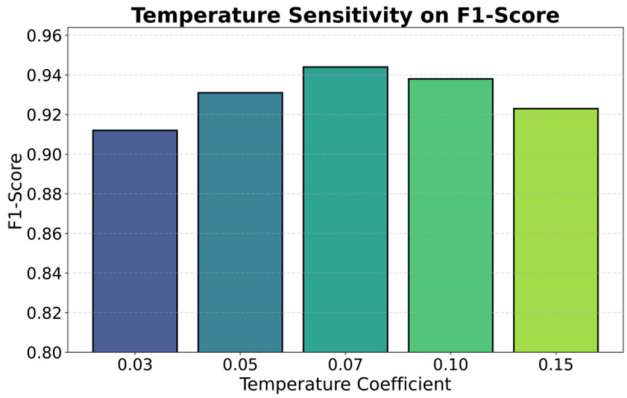

Figure 3. Experiment comparing the sensitivity of the temperature coefficient of loss to the F1-Score index.

From the overall trend, the F1 Score shows a steady increase followed by a slight decline as the temperature coefficient changes. This indicates that the temperature parameter in contrastive learning has a substantial impact on the distribution structure of the embedding space. When the temperature is low, similarities between features are amplified. The model tends to form overly concentrated embeddings when constructing positive and negative pairs. This limits the separability between abnormal and normal patterns, and the F1 Score remains at a relatively low level. As the temperature increases, the model is able to enlarge the distance between different semantic regions more effectively. This sharpens the classification boundary and leads to the optimal F1 Score.

In summary, the experiment shows that selecting an appropriate temperature coefficient is crucial for maintaining stability and discriminability of the embedding space across nodes in a federated contrastive learning framework. A moderate temperature strengthens semantic structures, improves cross-node consistency, and enhances the model's ability to capture rare abnormal patterns. Extremely high or low temperatures can damage the optimal embedding structure. A suitable temperature improves representation quality in federated settings and enables accurate and stable anomaly detection under privacy constraints.

5. Conclusions

This study proposes a federated contrastive learning based distributed representation framework to address privacy protection requirements in IoT anomaly detection. The method does not share raw data. It builds a unified and highly discriminative latent representation space through local encoding, cross-node consistency, and global aggregation. It enables accurate identification of abnormal patterns under multi-source heterogeneous data. Experimental results show that the framework maintains stable performance under complex distribution shifts and can effectively

 doi:10.20944/preprints202602.1749.v1

handle device diversity, communication constraints, and data security limitations in IoT systems. Overall, the method provides a system-level solution for secure and efficient IoT monitoring and establishes a solid foundation for future research and deployment.

In terms of model design, this work explores the feasibility and advantages of applying contrastive learning in federated settings. The framework strengthens the robustness of representations through temperature adjustment, embedding alignment, and weighted aggregation. It is suitable not only for network traffic-based anomaly detection but also for broader tasks in industrial IoT, smart city sensing, and intelligent medical device monitoring. The model constructs a shared semantic space across nodes while ensuring that data remains on the local device. It offers a technical approach to address the long-standing conflict between distribution inconsistency and privacy protection. As IoT networks continue to expand, the distributed scalability of this method will become increasingly valuable.

In addition, the study highlights the significant impact of data heterogeneity and structural dynamics in federated environments. The framework improves the global model's generalization ability through adaptive mechanisms. It automatically balances the contribution of different nodes and reduces the influence of noisy nodes. This enhances its engineering practicality in real-world applications. The framework can maintain high reliability during long-term operation and supports the development of intelligent and autonomous monitoring systems. It shows potential in security alerting, fault prediction, and critical infrastructure protection, and can contribute to improving IoT security governance.

Future work may evaluate the model on a larger scale and more complex federated environments and explore deeper integration with generative modeling and causal representation learning to improve the interpretability of abnormal mechanisms. Efficient communication compression strategies, lightweight architectures, and personalized federated learning methods can also be incorporated to support deployment on highly constrained edge devices. With ongoing progress in privacy computing and encrypted communication, the proposed federated contrastive learning framework has the potential to become a key component in next-generation large-scale intelligent IoT systems and to support secure, efficient, and trustworthy distributed sensing.

References

1. N. Wang, S. Shi, Y. Chen et al., "Feco: Boosting intrusion detection capability in IoT networks via contrastive learning," IEEE Transactions on Dependable and Secure Computing, 2025.
2. W. Zhai, F. Wang, L. Liu et al., "Federated semi-supervised and semi-asynchronous learning for anomaly detection in IoT networks," arXiv preprint arXiv:2308.11981, 2023.
3. M. Devi, P. Nandal and H. Sehrawat, "Federated learning-enabled lightweight intrusion detection system for wireless sensor networks: A cybersecurity approach against DDoS attacks in smart city environments," Intelligent Systems with Applications, p. 200553, 2025.
4. R. Morshedi and S. M. Matinkhah, "A comprehensive review of deep learning techniques for anomaly detection in IoT networks: Methods, challenges, and datasets," Engineering Reports, vol. 7, no. 9, p. e70415, 2025.
5. Z. Zhang, Y. Xue, H. Zhu, S. Li, Z. Wang and Y. Xiao, “CondenseGraph: Communication-efficient distributed GNN training via on-the-fly graph condensation,” arXiv preprint arXiv:2601.17774, 2026.
6. X. Yang, Y. Ni, Y. Tang, Z. Qiu, C. Wang and T. Yuan, “Graph-structured deep learning framework for multi-task contention identification with high-dimensional metrics,” arXiv preprint arXiv:2601.20389, 2026.
7. R. Ying, Q. Liu, Y. Wang and Y. Xiao, “AI-based causal reasoning over knowledge graphs for data-driven and intervention-oriented enterprise performance analysis,” 2025.
8. T. Guan, “A multi-agent coding assistant for cloud-native development: From requirements to deployable microservices,” 2025.
9. Y. Ou, S. Huang, R. Yan, K. Zhou, Y. Shu and Y. Huang, “A residual-regulated machine learning method for non-stationary time series forecasting using second-order differencing,” 2025.

 doi:10.20944/preprints202602.1749.v1

10. Y. Ni, X. Yang, Y. Tang, Z. Qiu, C. Wang and T. Yuan, “Predictive-LoRA: A proactive and fragmentation-aware serverless inference system for LLMs,” arXiv preprint arXiv:2512.20210, 2025.
11. R. Liu, L. Yang, R. Zhang and S. Wang, “Generative modeling of human-computer interfaces with diffusion processes and conditional control,” arXiv preprint arXiv:2601.06823, 2026.
12. Y. Wang, “Intelligent compliance risk detection in the pharmaceutical industry via transformer-driven semantic discrimination,” Transactions on Computational and Scientific Methods, vol. 4, no. 7, 2024.
13. A. Xie, “Adaptive privacy-aware federated language modeling for collaborative electronic medical record analysis,” Transactions on Computational and Scientific Methods, vol. 4, no. 8, 2024.
14. Z. Qiu, “A multi-scale deep learning and uncertainty estimation framework for comprehensive anomaly detection in cloud environments,” Transactions on Computational and Scientific Methods, vol. 3, no. 2, 2023.
15. H. Wang, C. Nie and C. Chiang, “Attention-driven deep learning framework for intelligent anomaly detection in ETL processes,” 2025.
16. Y. Hu, J. Li, K. Gao, Z. Zhang, H. Zhu and X. Yan, “TrustOrch: A dynamic trust-aware orchestration framework for adversarially robust multi-agent collaboration,” 2025.
17. Y. Xing, M. Wang, Y. Deng, H. Liu and Y. Zi, “Explainable representation learning in large language models for fine-grained sentiment and opinion classification,” 2025.
18. Y. Zhou, “A unified reinforcement learning framework for dynamic user profiling and predictive recommendation,” SSRN 5841223, 2025.
19. B. Chen, “FlashServe: Cost-efficient serverless inference scheduling for large language models via tiered memory management and predictive autoscaling,” 2025.
20. C. Zhang, C. Shao, J. Jiang, Y. Ni and X. Sun, “Graph-transformer reconstruction learning for unsupervised anomaly detection in dependency-coupled systems,” 2025.
21. S. Sun, “CIRR: Causal-invariant retrieval-augmented recommendation with faithful explanations under distribution shift,” arXiv preprint arXiv:2512.18683, 2025.
22. Y. Yin, J. Jang-Jaccard, F. Sabrina et al., "Improving multilayer-perceptron (MLP)-based network anomaly detection with BIRCH clustering on CICIDS-2017 dataset," Proceedings of the 2023 26th International Conference on Computer Supported Cooperative Work in Design (CSCWD), pp. 423-431, 2023.
23. G. Altangerel, M. Tejfel and E. Tsogbaatar, "IoT anomaly detection with 1D CNN using P4 capabilities," Acta Electrotechnica et Informatica, vol. 23, no. 2, pp. 3-12, 2023.
24. Z. Sun, A. M. H. Teixeira and S. Toor, "GNN-IDS: Graph neural network based intrusion detection system," Proceedings of the 19th International Conference on Availability, Reliability and Security, pp. 1-12, 2024.
25. C. Ding, S. Sun and J. Zhao, "MST-GAT: A multimodal spatial-temporal graph attention network for time series anomaly detection," Information Fusion, vol. 89, pp. 527-536, 2023.