Journal of Computer Technology and Software

ISSN:2998-2383

Vol. 3, No. 4, 2024

Secure Parameter Aggregation and Distributed Anomaly Detection in Cross-Cloud Data Center Networks

Shirui Zhang[1*]
[1]Northeastern University, Seattle, USA
*Corresponding author: Shirui Zhang; zsr1998zsr@hotmail.com

Abstract: This study proposes an intelligent method based on federated learning to address the problem of anomaly detection in distributed cross-cloud data center environments. The research first analyzes the complexity and risks in multi-tenant shared computing scenarios and points out the limitations of traditional centralized detection methods, including high bandwidth consumption, privacy leakage risks, and poor adaptability to heterogeneous data in large-scale distributed architectures. To solve these issues, cross-cloud data centers are modeled as a collaborative system composed of multiple autonomous nodes. Each node performs local data preprocessing and feature modeling, while secure parameter sharing and global aggregation are achieved through a federated learning framework. In method design, the framework integrates multi-layer representation encoding with local optimization, enabling the model to capture complex cross-tenant interactions without exposing raw data. A global weighted aggregation strategy further improves robustness and generalization under heterogeneous data distributions. The anomaly detection task is formulated as a joint optimization of latent feature representation and reconstruction error, and precise anomaly identification is achieved through global parameter updates. Overall analysis shows that the proposed method balances privacy protection and detection accuracy, alleviates communication burdens in cross-data center collaboration, and provides reliable support for the secure operation of critical systems across multiple industries.

Keywords: Federated learning; distributed anomaly detection; cross-cloud data center; privacy protection

1. Introduction

In the continuous development of the digital economy, cloud computing has become a fundamental infrastructure supporting social operations and industrial upgrading. As enterprises and institutions gradually migrate their core businesses to cloud platforms, the scale and complexity of distributed resource management and data interaction have sharply increased. The multi-tenant model of cross-cloud data centers improves efficiency and resource utilization but also brings risks to security and stability [1]. Abnormal behaviors may occur in computing nodes, communication links, or tenant interactions. If not identified in time, they can easily lead to service degradation, data leakage, or even system collapse. Under this background, how to perform efficient and precise anomaly detection in large-scale distributed environments has become a key scientific and engineering challenge [2].

Traditional centralized anomaly detection methods often fail in cross-cloud environments. First, the data volume in distributed systems grows exponentially. Transmitting all data to a single node for centralized analysis imposes huge bandwidth and storage pressure and may cause latency and bottlenecks. Second, strict privacy and compliance requirements among cloud service providers restrict the sharing of raw data across data centers. Centralized detection methods are, therefore, limited. In this case, there is a pressing need for a new paradigm that can ensure privacy while enabling collaborative detection in distributed environments [3].

The rise of federated learning provides a promising solution. It allows multiple cloud data centers to collaborate by sharing model parameters without exchanging raw data. This mechanism not only reduces the risk of data leakage but also preserves local data distribution while improving the generalization of the overall detection model. In cross-cloud scenarios, federated learning can overcome physical and privacy barriers by connecting decentralized computing and storage resources into a unified detection framework. This approach revitalizes distributed anomaly detection and offers a feasible path for collaborative security defense across institutions [4].

From an application perspective, distributed anomaly detection has broad significance beyond technical advances. It is essential to the digital transformation of finance, healthcare, energy, and manufacturing. In financial transaction systems, cross-data-center federated detection helps identify abnormal trades and fraud in real time [5]. In smart healthcare, hospital data centers can work together to detect attacks and service anomalies, ensuring patient data security and system reliability. In industrial internet systems, anomaly detection among multi-source sensors and control nodes supports stable equipment operation and continuous production processes. Thus, federated anomaly detection across clouds has both theoretical value and practical importance for building intelligent and secure industries.

In summary, studying federated learning-based cross-cloud distributed anomaly detection is both a timely response to new cloud security challenges and an inevitable step toward integrating intelligent computing with privacy protection. This research direction can build more efficient, reliable, and secure

detection systems in a world of increasingly interconnected data resources. It meets the practical needs of cross-cloud collaboration while aligning with the governance and compliance requirements of the digital society. Therefore, exploring this issue carries high academic value and directly concerns the robustness of information infrastructure and the sustainable growth of the digital economy.

2. Related work

In the field of distributed anomaly detection, much research has focused on traditional approaches. Early methods often relied on rule matching or statistical modeling, where thresholds or probability distributions were set to identify abnormal behaviors [6]. These methods are simple to implement and provide good interpretability, but they often fail in dynamic multi-tenant environments. On the one hand, monitoring based on a single feature dimension cannot capture complex multi-dimensional dependencies. On the other hand, as data centers expand and workloads fluctuate, static rules cannot adapt flexibly to new anomaly patterns. These limitations highlight the weaknesses of centralized detection mechanisms in cross-cloud environments.

With the development of machine learning and deep learning, model-based anomaly detection has gradually become mainstream. Researchers have built supervised or unsupervised frameworks, using multilayer neural networks to extract latent patterns from high-dimensional data, thereby improving accuracy and robustness. At this stage, most methods still rely on centralized data storage and training, but such approaches are difficult to apply at scale in cross-cloud environments. Large-scale data transmission across centers leads to heavy consumption of bandwidth and storage. At the same time, sensitive data cannot be freely shared across physical boundaries, and privacy and compliance requirements remain major obstacles. As a result, although traditional deep learning methods bring performance improvements, they cannot be well adapted to cross-cloud distributed scenarios [7].

To address the problems of data silos and privacy constraints, federated learning has entered the research landscape of anomaly detection. This method trains models locally on data and only shares parameter updates, enabling collaborative modeling across centers. Existing studies show that this mechanism can effectively avoid the risk of data leakage while maintaining consistency and coordination of detection models among data centers. In cross-cloud scenarios, federated learning not only resolves the difficulty of centralized data storage but also reduces bandwidth consumption and latency to some extent. However, current studies still face limitations. Heterogeneous data distributions and differences in communication efficiency may degrade global model performance. How to ensure both efficiency and accuracy while protecting privacy remains a critical challenge.

In recent years, cross-cloud distributed anomaly detection has shown a trend of multi-directional integration. On the one hand, research has increasingly introduced new modeling methods such as graph neural networks and attention mechanisms into federated learning frameworks, to better capture complex interactions and global dependencies. On the other hand, some studies have explored mechanisms such as communication compression, parameter weighting, and adaptive aggregation, aiming to alleviate efficiency bottlenecks in large-scale federated training. These advances provide new ideas and technical resources for collaborative anomaly detection in cross-cloud data centers, and they indicate that the field is rapidly evolving. Overall, existing research has laid a solid foundation for this topic, but improvements are still needed in robustness, generalization, and cross-platform adaptability. These gaps create broad opportunities for future studies.

Recent studies on cloud resource management and intelligent modeling provide additional support for cross-cloud anomaly detection. Constraint-aware virtual machine placement and queue-time prediction show that workload imbalance, resource matching, and service latency should be jointly considered when evaluating abnormal cloud states [8], [9]. Multi-view behavior modeling and structure-aware transformer modeling further demonstrate the value of combining heterogeneous resource traces with global structural context for detecting evolving anomalies [10], [11]. Hierarchical communication and computation scheduling for federated learning, together with temporal-context and entity-relation anomaly modeling, also highlight the need to balance local computation, aggregation efficiency, and relational risk representation in distributed intelligent systems [12], [13]. These studies strengthen the motivation for a federated cross-cloud detector that integrates privacy-preserving parameter sharing with global feature aggregation and adaptive participation control.

3. Proposed Approach

The core idea of this research is to implement distributed anomaly detection modeling across cloud data centers through a federated learning framework. In this cross-cloud scenario, each data center retains local data, completes model training locally, and periodically exchanges parameters with the global server. The overall model architecture is shown in Figure 1.

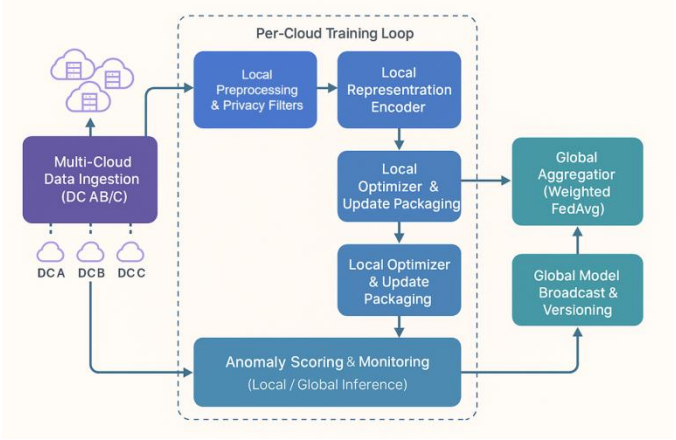

Figure 1. Overall model architecture

Assume that the data distribution of each data center is A, the corresponding number of local samples is B, and the total number of global samples is C. In the global optimization process, the goal is to minimize the weighted loss function of all data centers:

$$\min F(w) = \sum_{i=1}^{K} \frac{n_i}{N} F_i(w) \quad (1)$$

Where $F_i(w)$ represents the local loss function at the i-th data center, and w represents the global model parameters. This formula ensures that each data center contributes weight to the global optimization according to its data size, thereby achieving fair and effective collaborative training.

During the local training phase, each data center updates the model by gradient descent. For the i-th center, the parameter update formula for the t-th iteration is:

$$w_i^{t+1} = w_t - \eta \nabla F_i(w^t) \quad (2)$$

Where η is the learning rate, and $\nabla F_i(w^t)$ represents the gradient estimate based on the current global parameters at the i-th center. This mechanism ensures the independence of local computations, avoids cross-domain transmission of raw data, and enables each center to effectively optimize based on its own feature distribution.

In the parameter aggregation phase, all data centers upload their updated model weights or gradients to the global server, which performs weighted averaging to generate a new global model:

$$w_{t+1} = \sum_{i=1}^{K} \frac{n_i}{N} w_i^{t+1} \quad (3)$$

This aggregation strategy ensures that the contribution of different data centers to the global model matches the size of their data, thereby achieving performance improvements while ensuring privacy. To further enhance the robustness of the model, regularization constraints can be introduced during the aggregation process to limit excessive parameter drift.

At the anomaly detection modeling level, this study models the input time series data or interaction behavior as a high-dimensional feature vector $x \in R^d$ and maps it to the latent space through the deep representation function $\phi(\cdot)$:

$$h = \phi(x; w) \quad (4)$$

Where h is the potential representation. In the detection process, the anomaly score can be given by the reconstruction error or probability estimation function, such as when based on the reconstruction mechanism:

$$s(x) = \| x - \hat{x} \|_2^2 \quad (5)$$

Where $\hat{x}$ is the model's reconstruction of the input. By jointly optimizing the reconstruction error and the global aggregation objective within a federated learning framework, we can achieve accurate identification and dynamic adaptation of abnormal behavior in cross-cloud scenarios.

4. Performance Evaluation

4.1 Dataset

This study selects the CloudSim-based Data Center Workload Dataset as the main data source. The dataset consists of multi-tenant task scheduling and resource usage logs in typical cloud computing environments. It covers multiple dimensions such as virtual machine creation, task submission, CPU and memory utilization, bandwidth usage, and task execution delay. The data originates from simulated real cloud platform operations. It effectively reflects resource competition and interaction among tenants, providing rich contextual information for anomaly detection.

The dataset has strong temporal characteristics and diversity. It contains tens of thousands of task-level records. Each record corresponds to tenant activities and system resource states at different time slices. By comparing the differences between normal and abnormal tasks, patterns such as abnormal resource consumption, excessive task delay, and potential attack behaviors can be revealed. Since it includes different tenants and task types, the dataset is highly representative in reflecting the complexity of multi-tenant shared resource environments.

In addition, the dataset also has transferability in cross-cloud scenarios. By dividing subsets of data centers or tenants, it can simulate distributed operating environments across platforms. This enables the study to maintain data consistency while exploring the applicability of distributed anomaly detection methods under different resource conditions and system configurations. Analysis and modeling based on this dataset help verify the feasibility and robustness of cross-cloud anomaly detection frameworks across multiple dimensions and scenarios.

4.2 Experimental Results

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

Table 1. Comparative experimental results

Method	AUC	ACC	F1-SCORE	Recall
BILSTM [14]	0.923	0.901	0.892	0.879
1DCNN [15]	0.936	0.914	0.905	0.892
ConvNextv2 [16]	0.948	0.922	0.913	0.906
CNN+Transformer [17]	0.957	0.931	0.922	0.917
Ours	0.971	0.944	0.937	0.929

From the overall results, all methods achieve high levels of AUC, ACC, F1-Score, and Recall. This indicates that deep learning frameworks generally have strong modeling capabilities in anomaly detection tasks across cloud data centers. However, the differences among methods also highlight the gap in their ability to capture multi-dimensional dependencies and global features. BiLSTM, as a traditional sequence modeling method, has certain advantages in modeling long-term dependencies. Yet, when facing complex multi-tenant interactions in cross-cloud environments, its feature extraction remains insufficient. As a result, it shows the lowest performance across all metrics.

With the introduction of 1DCNN, the model can extract local features through convolutional structures. Its performance improves significantly on all four metrics, especially in accuracy and F1-Score, where it shows stronger robustness. ConvNeXtV2 further enhances the depth and hierarchy of feature extraction. It performs better in modeling nonlinear features and complex distributions. Therefore, it outperforms the first two methods in AUC and Recall. This demonstrates

that convolutional structures are adaptable in high-dimensional data modeling, but their ability to capture global dependencies in cross-cloud environments is still limited.

The combination of CNN and Transformer brings further breakthroughs in performance. Convolutional networks capture local patterns, while Transformers are strong in modeling long-range dependencies. This approach balances local details and global features. It improves performance on all four metrics compared with the previous method. The improvement in Recall is especially notable, indicating that this model is more sensitive in identifying potential anomalies. This makes it more suitable for anomaly detection in distributed environments.

In comparison, the method proposed in this study achieves the highest results on all four metrics. Its advantages are most prominent in AUC and F1-Score. This shows that the proposed model can better balance local and global feature modeling in cross-cloud scenarios. At the same time, with the support of privacy protection and distributed optimization mechanisms, it demonstrates stronger generalization and stability. Therefore, the method not only has theoretical innovation but also shows great potential in practice. It provides a more efficient solution for secure and reliable operations in cross-cloud data centers.

This paper further gives the impact of the federated client learning rate on the experimental results, and the experimental results are shown in Figure 2.

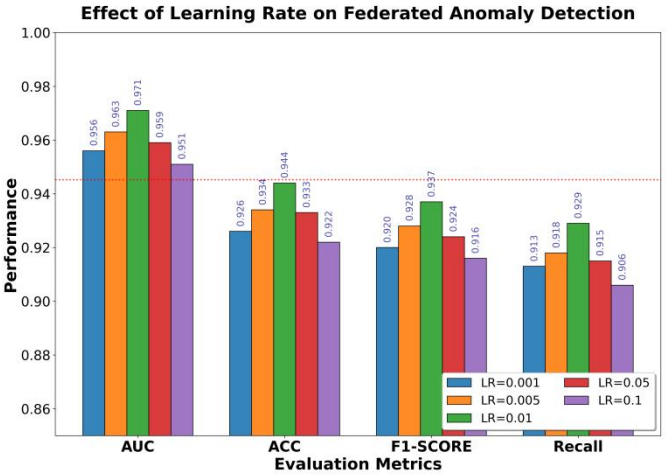

Figure 2. The impact of federated client learning rate on experimental results

From the overall results, different learning rate settings have a significant impact on model performance. The AUC metric reaches its highest value when the learning rate is 0.01, indicating that a moderate learning rate can better balance parameter update speed and gradient convergence. When the learning rate is too low, the convergence during global aggregation is slow, which reduces the ability of feature representation. When the learning rate is too high, updates are faster but introduce large fluctuations, which decreases stability and generalization in detection.

For the accuracy metric, the learning rate of 0.01 also shows the best performance, with a clear gap compared to both lower and higher learning rates. This suggests that a reasonable learning rate not only helps the model reach convergence faster but also avoids performance degradation caused by overfitting or underfitting. In particular, under cross-cloud environments with highly heterogeneous data distributions, a stable learning rate helps the global model maintain consistency across multiple tenants and tasks.

From the F1-Score results, the learning rate has a strong influence on the sensitivity of anomaly detection. When the learning rate is within the optimal range, the model can maintain a good balance between precision and recall, leading to more stable overall detection. When the learning rate is high, the F1-Score decreases significantly, showing that the model fluctuates when capturing anomaly patterns. This may cause some anomalies to be missed or misclassified, which brings potential risks in practical cross-cloud scenarios.

Finally, the recall metric shows a similar trend. A moderate learning rate maximizes the sensitivity of the model to anomalies and improves the detection rate. This is especially important for anomaly detection in cross-cloud data centers, as missed detections may cause serious security and resource scheduling problems. Overall, the experimental results confirm the critical role of learning rate in federated learning optimization. They highlight the importance of reasonable hyperparameter selection in ensuring accuracy and robustness of detection in cross-cloud distributed environments.

This paper also gives the impact of the proportion of participating clients on the experimental results, and the experimental results are shown in Figure 3.

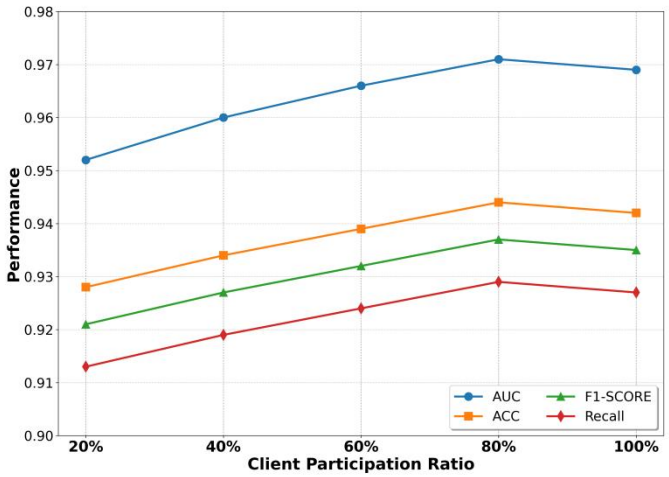

Figure 3. The impact of the proportion of participating clients on experimental results

From the overall trend, it can be seen that as the client participation ratio increases, the performance of the model on all metrics shows a gradual improvement. The rise in AUC is the most obvious. It reaches its peak when the participation ratio is 80 percent. This indicates that the involvement of more clients enhances the model's ability to capture global feature distributions. In cross-cloud data center environments, different clients often correspond to heterogeneous data sources. A higher participation ratio improves the generalization ability of the global model, enabling it to capture anomaly patterns more comprehensively.

For accuracy, ACC steadily increases as the participation ratio grows and reaches its best level at around 80 percent. This result shows that the proportion of participating clients directly affects the discriminative power of the model. A higher participation ratio helps to reduce the bias caused by data

imbalance. In distributed scenarios, each client has limited data. A low participation ratio may cause insufficient information during training, which limits the discriminative performance of the global model.

The variation of F1-Score is also closely related to the participation ratio. It achieves the best performance at 80 percent and then drops slightly at 100 percent. This shows that in cross-cloud environments, more clients are not always better. Too high a participation ratio may bring communication overhead and unstable model updates. These factors may interfere with the balance between precision and recall. This result suggests that in practical deployment, it is necessary to find a reasonable trade-off between model performance and system cost to ensure efficient and stable detection.

For recall, the results show a similar upward trend to the other metrics. This indicates that a higher client participation ratio improves the model's sensitivity and detection ability for anomalies. This is critical for anomaly detection in cross-cloud data centers, since missed detections can lead to severe system risks. Overall, the experimental results demonstrate that the client participation ratio has an important impact on distributed anomaly detection under federated learning. A reasonable choice of participation ratio not only enhances model performance but also enables more efficient resource utilization and stronger security assurance in cross-cloud scenarios.

5. Conclusion

This study focuses on the problem of distributed anomaly detection in cross-cloud data center environments and proposes a solution based on federated learning. The research starts from the complexity of multi-tenant shared computing resources and considers the dual demands of data privacy and system security in cross-cloud scenarios. A detection framework is constructed that balances local modeling and global collaboration. By sharing parameters without transmitting raw data, the method effectively alleviates the problems of data silos and privacy leakage. At the same time, it enables the recognition of complex anomaly patterns on a global scale. This approach not only enriches the theoretical system of anomaly detection but also provides new insights for privacy protection in distributed intelligent computing.

From the overall analysis, this study highlights the necessity of cross-cloud collaboration and the feasibility of federated optimization. Against the backdrop of traditional centralized methods being unsuitable for distributed and heterogeneous environments, the proposed method shows significant advantages. It balances local data features and global model consistency. It also mitigates challenges caused by uneven data distribution and differences in computing resources. Therefore, the results of this research not only extend the application boundaries of anomaly detection but also provide important reference values for resource scheduling and security governance in cross-cloud architectures.

At the application level, the proposed method has broad implications for finance, healthcare, industrial manufacturing, and energy. As these industries increasingly rely on collaboration and sharing across data centers, the accuracy and timeliness of anomaly detection directly affect business security and reliability. The federated detection framework proposed in this study improves detection performance while protecting data privacy. It is expected to provide stable security assurance for various types of critical infrastructure. Its contribution to industry practice is reflected in enhancing detection efficiency, reducing risk costs, and strengthening overall system resilience. These advantages show its potential for real-world deployment.

Future research can further expand and deepen the contributions of this study. On the one hand, more efficient communication and aggregation strategies can be explored to reduce latency and overhead in large-scale federated training. On the other hand, multimodal data and adaptive optimization mechanisms can be combined to enhance the ability to capture complex anomaly patterns in cross-domain environments. In addition, under the trend of increasing inter-organizational collaboration and compliance requirements, the proposed framework can provide theoretical and technical support for building a more open, secure, and trustworthy digital ecosystem. These directions will not only advance academic research but also have a profound impact on the security and intelligence of practical application scenarios.

References

[1] W. Chen, Y. Lu, B. Li, et al., "FedDetect: Federated Learning for Anomaly Detection in Industrial Internet of Things," IEEE Transactions on Industrial Informatics, vol. 19, no. 5, pp. 6034-6045, 2023.

[2] Y. J. Liu, G. Feng, D. Niyato, et al., "Ensemble distillation based adaptive quantization for supporting federated learning in wireless networks," IEEE Transactions on Wireless Communications, vol. 22, no. 6, pp. 4013-4027, 2022.

[3] M. Mahmoudi, A. Erbad, K. Qaraqe, et al., "Federated Learning for Anomaly Detection: A Survey," IEEE Access, vol. 11, pp. 104698-104721, 2023.

[4] W. Jianping, Q. Guangqiu, W. Chunming, et al., "Federated learning for network attack detection using attention-based graph neural networks," Scientific Reports, vol. 14, no. 1, p. 19088, 2024.

[5] S. Gayathri and D. Surendran, "Unified ensemble federated learning with cloud computing for online anomaly detection in energy-efficient wireless sensor networks," Journal of cloud computing, vol. 13, no. 1, p. 49, 2024.

[6] H. Liu, Z. Chen, Z. Zhang, et al., "Federated Learning for Anomaly Detection in Distributed Cyber-Physical Systems," IEEE Internet of Things Journal, vol. 10, no. 8, pp. 6918-6930, 2023.

[7] S. Ma, J. Nie, J. Kang, et al., "Privacy-preserving anomaly detection in cloud manufacturing via federated transformer," IEEE Transactions on Industrial Informatics, vol. 18, no. 12, pp. 8977-8987, 2022.

[8] J. Kou and E. Pei, "Constraint-Aware Resource Matching for Virtual Machine Placement in Cloud Environments," 2024.

[9] H. Zhuang and K. Gao, "Global Context Modeling and Structure-Guided Feature Enhancement for Queue Time Prediction in Large-Scale Cloud Systems," 2024.

[10] X. Zhang and Z. Fang, "Deep Learning-Based Multi-View Behavioral Modeling and Trend Regression for Cloud Resource Prediction," 2024.

[11] K. Zeng, "Transformer-Based Structure-Aware Lung Cancer Image Segmentation with Multi-Scale Fusion and Boundary-Guided Prediction," 2024.

[12] R. Meng, "Hierarchical Communication and Computation Scheduling for Efficient Federated Learning in Cloud-Edge Collaborative Intelligent Systems," 2024.

[13] R. Yan and M. Guo, "Enterprise Audit Risk Identification Through Temporal Context and Entity Relationship Anomaly Modeling," 2024.

[14] J. Zhang, X. Zhang, Z. Liu, et al., "A network intrusion detection model based on BiLSTM with multi-head attention mechanism," Electronics, vol. 12, no. 19, p. 4170, 2023.

[15] E. U. H. Qazi, A. Almorjan, and T. Zia, "A one-dimensional convolutional neural network (1D-CNN) based deep learning system for network intrusion detection," Applied Sciences, vol. 12, no. 16, p. 7986, 2022.

[16] V. S. Balusa and K. Srinivas, "An effective intrusion detection system for edge computing using ConvNeXt and ResNet152V2," International Journal of Computational Intelligence and Applications, vol. 23, no. 03, p. 2450014, 2024.

[17] H. Kamal and M. Mashaly, "Advanced hybrid transformer-CNN deep learning model for effective intrusion detection systems with class imbalance mitigation using resampling techniques," Future Internet, vol. 16, no. 12, p. 481, 2024.