Artificial Intelligence and Computing Innovations

Intelligence Science Publishing

https://www.ispublishing.com

ISSN: 3070-4944

Vo.4, No.2, 2024

Enhancing Cloud Security with Graph Neural Network-Driven Anomaly Detection in Multi-Tenant Environments

Saina Shawulieti[1]*

[1]*University of California, San Diego, La Jolla, USA*

**Corresponding author: sainashawulieti@gmail.com*

Abstract: This paper addresses the key problem of anomaly detection in multi-tenant cloud environments and proposes an intelligent detection method based on graph neural networks. The study first analyzes the complexity and security risks faced during the process of resource sharing among multiple tenants. It points out that traditional methods based on rules or single features fail to accurately capture tenant interactions and potential abnormal patterns. On this basis, the multi-tenant environment is abstracted into a graph structure, where nodes represent tenants or resources and edges represent interactions or dependencies. Graph neural networks are then used to take advantage of information propagation and feature aggregation, enabling global modeling and representation of tenant behaviors. The method combines graph convolution and attention mechanisms, allowing the model to dynamically focus on key neighbor nodes within multi-level dependencies. This enhances sensitivity and discriminative ability toward abnormal behaviors. At the same time, a joint loss function is constructed to balance feature reconstruction and regularization constraints. This design effectively strengthens the generalization ability and robustness of the model. To verify the effectiveness of the proposed method, systematic experiments were conducted on public cloud computing datasets. Sensitivity analyses were also performed on hyperparameters and environmental conditions. The results show that the method outperforms baseline models in key metrics such as AUC, ACC, F1-Score, and Precision. It maintains stable performance under different conditions. These findings confirm that the proposed anomaly detection framework with graph neural networks has strong practical value and application potential in complex multi-tenant cloud scenarios.

Keywords: Multi-tenant cloud environment; abnormal behavior recognition; graph neural network; robustness

1. Introduction

Cloud computing, as the core form of next-generation information infrastructure, has become a key pillar supporting the digital economy and intelligent society. Through technologies such as virtualization, containerization, and elastic scheduling, it provides users with on-demand allocation, flexible scaling, and

high availability of computing and storage resources. In this model, multi-tenancy is an essential characteristic. Multiple users or applications share the same physical or virtual environment to improve resource utilization and cost efficiency[1]. However, with the continuous growth of users and the increasing complexity of business scenarios, tenant behavior in cloud platforms has become highly dynamic and diverse. This leads to frequent abnormal behaviors and rising security risks. Whether caused by malicious attacks resulting in unauthorized access or abnormal fluctuations in resource usage patterns, such issues undermine the stability and reliability of multi-tenant environments. Therefore, how to achieve efficient and intelligent anomaly detection in complex shared environments has become a pressing challenge for both academia and industry[2].

Traditional anomaly detection methods often rely on rule matching, statistical modeling, or supervised learning based on single features. These approaches can be effective in static environments but face severe limitations in multi-tenant cloud settings. On the one hand, tenant behavior is highly time-varying and correlated, and simple thresholds or single-dimension features fail to capture complex interaction patterns. On the other hand, abnormal behaviors are usually hidden within large volumes of normal operations. They are infrequent, imbalanced, and evolving. Traditional methods struggle to ensure accuracy while maintaining real-time performance and scalability. Moreover, as cloud platforms continue to expand, resource sharing and business dependencies among tenants grow increasingly complex. Analyzing data from a single point of view is no longer sufficient to reveal the underlying global structural features. Thus, integrating multi-source heterogeneous data and establishing global dependencies among tenants have become key breakthroughs for innovation in anomaly detection[3].

In recent years, graph structures have emerged as an important tool for representing complex system interactions. In multi-tenant cloud environments, users, virtual machines, containers, services, and resources naturally form multi-layered and multi-relational graphs. Nodes represent entities, and edges describe interactions and dependencies among tenants. This feature provides the basis for applying graph neural networks[4]. Unlike traditional methods, graph neural networks can propagate and aggregate information along graph topologies, thereby capturing higher-order associations and latent abnormal patterns. By modeling with graphs, it becomes possible to analyze tenant behaviors from a global perspective and to detect hidden abnormal paths and propagation effects within complex interaction networks. This lays the foundation for more precise and intelligent anomaly detection[5].

Introducing graph neural networks for anomaly detection in cloud environments has both theoretical and practical significance. It helps strengthen the security and reliability of cloud platforms by preventing service disruptions, data leaks, or resource abuse caused by anomalies. It also enhances user trust in cloud services[6]. At the same time, such methods provide intelligent support for operations in multi-tenant settings. They assist providers in resource scheduling, traffic optimization, and policy formulation by enabling more scientific risk identification. Furthermore, scalable anomaly detection frameworks can adapt to different cloud scales and types, offering solid support for the sustainable development of cloud computing.

In summary, anomaly detection in multi-tenant cloud environments has become a core issue for ensuring the quality and security of cloud services. Graph neural network-based methods can overcome the limitations of traditional approaches by exploiting complex relational information among tenants and forming a comprehensive analysis framework with multiple levels and perspectives[7]. This not only improves detection accuracy and robustness but also advances intelligent operations and security defense systems in cloud platforms. Research in this direction is of strategic importance and practical value for promoting the healthy development of cloud computing, building next-generation intelligent infrastructures, and safeguarding the secure operation of the digital economy.

Recent research on operational graph reasoning and backend observability further strengthens the technical foundation for graph-based anomaly detection in multi-tenant clouds. Time-varying operational graph reasoning and risk-constrained repair planning show that dynamic dependency graphs can connect

anomaly identification with downstream remediation decisions[8]. Behavior representation learning for AI inference backend monitoring indicates that runtime traces can be encoded into compact behavioral states for reliable system monitoring[9], while transformer-based semantic log modeling improves robust and explainable cloud security analytics by extracting higher-level semantics from operational logs[10]. Fine-grained backend anomaly classification based on structural dependency learning and version-aware feature alignment highlights the value of incorporating dependency structure and software evolution into detection features[11]. Graph neural modeling of traffic propagation and topological interactions provides further support for describing how anomalous activity spreads through cloud networks[12]. Reinforcement learning-driven dynamic cache management also suggests that hotspot-aware operational policies can be integrated with security monitoring to improve adaptive control under shifting workloads[13]. Together, these studies support the design of a GNN-based multi-tenant anomaly detection framework that combines structural representation, temporal context, semantic log evidence, and operational feedback.

2. Proposed Approach

In a multi-tenant cloud environment, the interactions between tenants, resources, and services can be naturally abstracted as a graph structure, where nodes represent different entities (such as virtual machines, containers, application services), and edges represent the dependencies or interactions between them. The overall model architecture is shown in Figure 1.

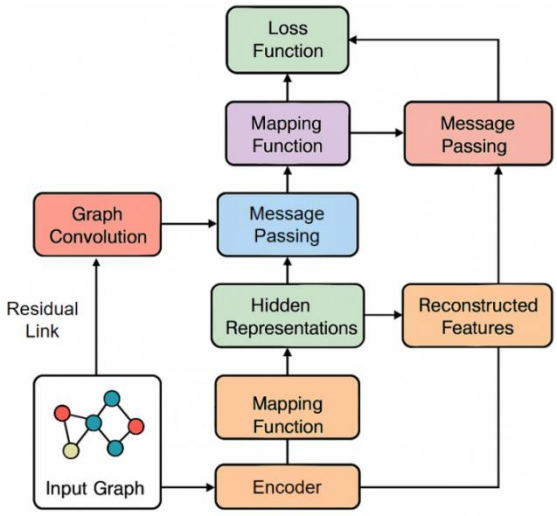


Figure 1. Overall model architecture

First, given a set of tenants V and an interaction set E, we can construct a graph $G=(V,E)$ of a multi-tenant cloud environment. The adjacency matrix of the graph can be defined as:

$$A_{ij}=1 \text{ if tenants } i \text{ and } j \text{ interact}; \quad A_{ij}=0 \text{ otherwise} \tag{1}$$

On this basis, each node v_i has an initial feature vector $x_i \in R^d$, which is used to describe its resource usage, access pattern, or context information. By combining node features into a matrix $X \in R^{|V|\times d}$, the original input can be represented as graph-structured data.

To capture the multi-order relationships between nodes in the graph, the message passing mechanism of the graph neural network is introduced. At the lth layer, the hidden representation of each node is updated by aggregating neighbor information and its own features. Its basic form is:

$$h_i^{(l+1)} = \sigma\left(\sum_{j \in N(i)} \frac{1}{c_{ij}} W^{(l)} h_j^{(l)}\right) \tag{2}$$

Where $h_j^{(l)}$ represents the representation of node i at layer l, $N(i)$ is the neighbor set, c_{ij} is the normalization coefficient, $W^{(l)}$ is the learnable weight matrix, and $\sigma(\cdot)$ is the nonlinear activation function. This process enables iterative updates of node representations, thereby gradually integrating interaction information between tenants.

Considering the heterogeneity and importance differences of interaction relationships in a multi-tenant cloud environment, a graph attention mechanism is introduced to improve the model's attention to key neighbors. For any two nodes i and j, their attention weight can be expressed as:

$$\alpha_{ij} = \frac{\exp\left(LeakyRELU\left(a^T [Wh_i \| Wh_j]\right)\right)}{\sum_{k \in N(i)} \exp\left(LeakyRELU\left(a^T [Wh_i \| Wh_k]\right)\right)} \tag{3}$$

Where $\|$ represents vector concatenation and a is the learnable parameter vector. Attention-based weighted aggregation can be further written as:

$$h_i^{(l+1)} = \sigma\left(\sum_{j \in N(i)} \alpha_{ij} W h_j^{(l)}\right) \tag{4}$$

This mechanism can dynamically adjust the contribution of neighbor nodes in representation learning, thereby enhancing sensitivity to abnormal behaviors.

After completing multi-layer information propagation and representation learning, a global discriminant model needs to be constructed to identify abnormal behavior. First, the final node representation obtained by the graph convolution or attention mechanism is denoted as $H \in R^{|V| \times d'}$. Through the mapping function $f(\cdot)$, it can be converted into an anomaly score vector $s \in R^{|V|}$, which is in the form of:

$$s_i = f(h_i) = \sigma(W_o h_i + b_o) \tag{5}$$

Where W_o and b_o are the output layer parameters. To optimize the anomaly detection model, a joint loss function based on reconstruction error and regularization is introduced:

$$L = \frac{1}{|V|} \sum_{i=1}^{|V|} \|x_i - \hat{x}_i\|_2^2 + \lambda \|W\|_F^2 \tag{6}$$

Where $\hat{x}_i$ is the reconstructed feature of the node, $\|\cdot\|_F^2$ is the Frobenius norm, and λ is the regularization weight. This loss function can maintain the model's ability to fit normal behavior while suppressing overfitting, ensuring robust recognition of abnormal behavior.

3. Performance Evaluation

3.1 Dataset

The dataset used in this study is the Azure Public Cloud Trace 2019. It is composed of operation logs from a real cloud platform and records of multi-tenant resource usage. The data includes events such as the creation, termination, and migration of virtual machines, as well as the usage of CPU, memory, and network resources. It also contains information about task dependencies among tenants and resource scheduling. With its large scale and long time span, the dataset fully reflects the dynamic and diverse characteristics of multi-tenant cloud environments.

The dataset has a rich structure that includes multiple types of records, such as resource usage time series, mappings between tenants and virtual machines, scheduling events, and system logs. These multimodal data provide a comprehensive view of tenant behavior patterns in shared environments and supply valuable context for anomaly detection. In particular, the cross-tenant interaction data and task dependency information offer a solid basis for building graph structures and modeling global correlations.

The reason for choosing this dataset lies in its advantages of real-world scenarios, multidimensional features, and rich abnormal patterns. It effectively supports the modeling requirements of graph neural networks. Compared with synthetic datasets or small-scale test sets, it is closer to practical production environments. This makes the proposed method more applicable and valuable for real-world deployment.

3.2 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	Precision
Transformer[14]	0.913	0.887	0.872	0.868
1DCNN[15]	0.901	0.876	0.861	0.857
FedFormer[16]	0.924	0.892	0.879	0.874
KAN[17]	0.932	0.899	0.886	0.882
Ours	0.948	0.914	0.903	0.899

From the results in Table 1, it can be observed that different methods show significant performance differences in anomaly detection tasks for multi-tenant cloud environments. 1DCNN, as a model that relies only on convolutional structures, has limitations in capturing global sequence dependencies and cross-tenant interactions. As a result, its overall performance is relatively low, with AUC, ACC, F1-Score, and Precision all falling behind the other methods. Transformer shows clear advantages over 1DCNN in global feature modeling, achieving improvements across the four metrics, but it still faces challenges in modeling temporal dynamics and complex dependencies.

FedFormer performs slightly better than Transformer. This indicates that the introduction of improved temporal modeling and feature fusion mechanisms can mitigate the information decay problem in long sequence modeling to some extent. The model can better distinguish between normal and abnormal behaviors. The improvement is particularly notable in ACC and F1-Score, reflecting its stronger ability to capture temporal and contextual relationships in multi-tenant environments. However,

FedFormer still fails to fully leverage tenant-to-tenant relational information, which limits its precision in fine-grained anomaly detection.

The KAN model further improves performance, with higher AUC and Precision values. This shows that it can more effectively separate abnormal and normal samples in the feature space. The method emphasizes structured modeling and nonlinear mapping capabilities, making it more robust than Transformer-based methods in complex cloud environments. Yet, due to the lack of deep modeling of graph-structured relations among tenants, KAN still has room for improvement in capturing hidden interaction patterns and potential anomaly propagation paths. Therefore, its overall performance has not reached the best level.

The proposed method achieves the best results across all metrics, with an AUC of 0.948, ACC reaching 0.914, F1-Score at 0.903, and Precision at 0.899. This demonstrates that introducing graph neural networks into multi-tenant anomaly detection can significantly enhance the model’ s ability to represent complex interactions and high-dimensional features. By constructing graph structures and integrating attention mechanisms, the method fuses global dependencies with local features and improves sensitivity to abnormal patterns. As a result, it achieves higher detection accuracy and robustness in real cloud computing scenarios.

This paper also gives an experiment on the sensitivity of the learning rate to AUC, and the experimental results are shown in Figure 2.

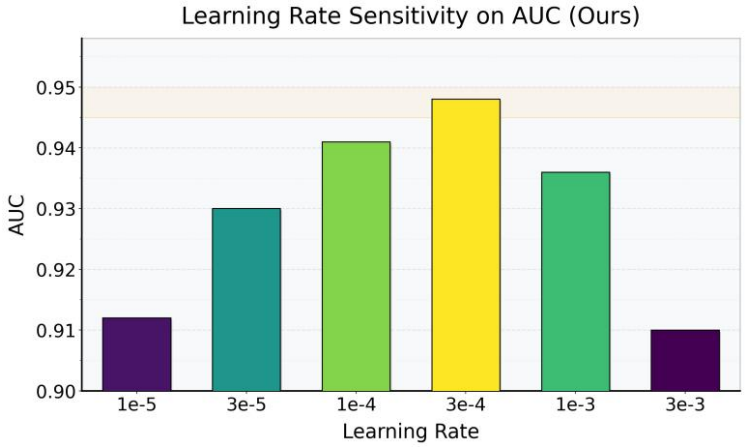


Figure 2. Experiment on the sensitivity of learning rate to AUC

From Figure 2, it can be seen that different learning rates have a significant impact on the AUC performance of the model in anomaly detection tasks for multi-tenant cloud environments. When the learning rate is too small, such as 1e-5 or 3e-5, the convergence speed of the model is limited. It fails to fully capture multi-level interactions in complex graph structures. As a result, the overall detection performance remains low, with AUC scores around 0.91, which is clearly lower than other ranges.

With a moderate increase in the learning rate, the model can learn tenant behavior patterns more quickly during gradient updates. It avoids overly smooth representations and achieves better performance in the 1e-4 and 3e-4 ranges. Among them, the learning rate of 3e-4 reaches the highest AUC value of 0.948. This indicates that such a setting strikes the best balance between effective convergence and stable optimization. It better captures multi-tenant interaction relationships and potential abnormal behavior patterns.

When the learning rate is further increased to 1e-3, the model still maintains some robustness, but performance begins to decline. This shows that an excessively large update step can cause gradient oscillation or even overfitting. The ability of the model to distinguish boundaries between abnormal and

normal patterns decreases. In highly dynamic and noisy cloud environments, fast parameter updates often amplify fluctuations in data distribution and reduce the stability of global modeling.

At a learning rate of 3e-3, model performance drops further, with AUC decreasing to a level close to that of very small learning rates. This result confirms the critical role of learning rate in the training process. Too small makes learning ineffective, while too large makes training unstable. Overall, a moderate learning rate range can best exploit the advantages of graph neural networks in anomaly detection for multi-tenant cloud environments. It ensures that the model has sufficient learning ability while maintaining robustness and generalization under complex interaction relationships.

This paper further presents the experimental results on the sensitivity of multi-tenant concurrency to experimental results, as shown in Figure 3.

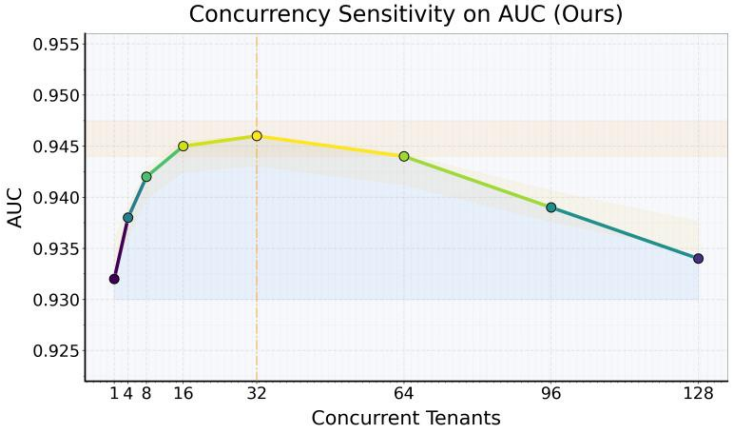


Figure 3. Experiment on the sensitivity of multi-tenant concurrency to experimental results

From the results in Figure 3, it can be seen that multi-tenant concurrency has a direct impact on the performance of anomaly detection models. When concurrency is low, such as 1, 4, or 8, the AUC values are relatively low. The main reason is the lack of sufficient tenant interactions. The global dependencies contained in the graph structure are limited. As a result, the model cannot fully learn complex behavior patterns and potential abnormal features, which restricts detection performance.

As concurrency increases to 16 and 32, the model performance improves, reaching a peak AUC of about 0.946 at 32 concurrent tenants. This shows that moderate concurrency enhances the diversity of tenant interactions and behavioral complexity. The graph neural network can capture more comprehensive dependencies during information propagation and representation learning. This significantly improves the ability to identify abnormal behaviors. It also indicates that multi-tenant interactions play an important role in strengthening feature representation in anomaly detection.

However, when concurrency continues to rise to 64 or more, the AUC begins to decline. This indicates that high concurrency provides richer interaction information but also introduces stronger noise and interference. This may create too many redundant edges or weakly related connections in the graph structure, reducing the ability of the model to focus on key abnormal patterns. In other words, excessive concurrency amplifies random fluctuations in complex environments, leading to reduced stability and robustness in detection results.

Overall, the results reveal the sensitivity of anomaly detection performance to concurrency. Too little concurrency fails to form effective global associations, while too much introduces excessive noise. Only within a moderate range can optimal performance be achieved. This conclusion not only provides a basis for model hyperparameter selection but also suggests that in real cloud computing scenarios, resource scheduling and tenant distribution should be reasonably controlled to maximize the performance and stability of anomaly detection systems.

4. Conclusion

This study addresses the problem of anomaly detection in multi-tenant cloud environments and proposes a modeling approach based on graph neural networks. By constructing a graph of tenants, resources, and interactions, the model captures complex dependency patterns from a global perspective and effectively integrates local and global features in node representations. The combination of graph convolution and attention mechanisms strengthens feature expression and enhances sensitivity to abnormal patterns. As a result, the model shows superiority in identifying potential risks in complex environments. The results confirm the importance of structured modeling in cloud computing and provide a new technical pathway for intelligent security protection in multi-tenant settings.

At the methodological level, this study emphasizes that anomaly detection in cloud computing should not rely only on single-point features or one-dimensional statistics. Instead, multi-level graph modeling is required to reflect interactions among tenants. Experimental analysis shows that proper parameter settings and environmental conditions are critical for model performance. For example, sensitivity experiments on learning rate and concurrency reveal the stability and adaptability of the model under different conditions. These findings demonstrate that the method has good controllability and can adapt to cloud platforms of different scales and complexities. This lays a solid foundation for broader applications.

In terms of application value, this study has strong practical significance. Multi-tenant cloud environments are the core infrastructure of the modern digital economy, and anomaly detection is directly related to service reliability and security. The proposed method can effectively identify potential risks, help cloud service providers improve operations, and reduce the likelihood of resource abuse and service disruptions. At the same time, the method is highly scalable and can be applied in financial, healthcare, and government clouds where strict requirements for security and stability exist. This supports the development of more intelligent and automated approaches to data security and resource management across the industry.

Looking forward, anomaly detection methods based on graph neural networks still have broad research potential. On the one hand, integrating causal reasoning and self-supervised learning can further enhance the detection of unknown anomalies. On the other hand, with the rise of edge-cloud collaboration and multimodal data, combining heterogeneous sources of information for real-time detection will be an important challenge. In addition, integrating this method with resource orchestration and policy optimization in cloud-native architectures could lead to the development of more intelligent closed-loop security systems. It can be expected that this study not only provides valuable insights for current anomaly detection in cloud computing but also offers important implications for building safer, smarter, and more efficient cloud ecosystems in the future.

5. References

[1] Liu Y, Kang Y, Xing C, et al. A Secure Federated Learning Framework for Anomaly Detection in Industrial Internet of Things[J]. IEEE Transactions on Industrial Informatics, 2023, 19(2): 1775-1785.

[2] Vinayakumar R, Alazab M, Soman K P, et al. Deep Learning Approach for Intelligent Intrusion Detection System[J]. IEEE Access, 2019, 7: 41525-41550.

[3] Marbel R, Cohen Y, Dubin R, et al. Cloudy with a Chance of Anomalies: Dynamic Graph Neural Network for Early Detection of Cloud Services' User Anomalies[J]. arXiv preprint arXiv:2409.12726, 2024.

[4] Zhao Z, Xiao Z, Tao J. MSDG: Multi-scale dynamic graph neural network for industrial time series anomaly detection[J]. Sensors, 2024, 24(22): 7218.

[5] Deng A, Hooi B. Graph neural network-based anomaly detection in multivariate time series[C]//Proceedings of the AAAI conference on artificial intelligence. 2021, 35(5): 4027-4035.

[6] Aceto G, Ciuonzo D, Montieri A, Pescapé A. Mobile Encrypted Traffic Classification Using Deep Learning: Experimental Evaluation, Lessons Learned, and Challenges[J]. IEEE Transactions on Network and Service Management, 2019, 16(2): 445-458.

[7] Zambon D, Alippi C, Livi L. Concept drift and anomaly detection in graph streams[J]. IEEE transactions on neural networks and learning systems, 2018, 29(11): 5592-5605.

[8] Tang Y. Time-Varying Operational Graph Reasoning and Risk-Constrained Repair Planning for Tool-Augmented Large Language Model Agents[J]. 2024, 4(6).

[9] Ding T, Yuan Y. Behavior Representation Learning for Reliable Monitoring of AI Inference Backends[J]. 2024, 4(1).

[10] Yang Z. Transformer-Based Semantic Log Modeling for Robust and Explainable Cloud Security Analytics[J]. 2024, 4(5).

[11] Liu W. Fine-Grained Backend Anomaly Classification through Structural Dependency Learning and Version-Aware Feature Alignment[J]. 2024, 4(4).

[12] Huang X. Capturing Traffic Propagation and Topological Interactions in Cloud Networks via Graph Neural Networks[J]. 2024, 4(5).

[13] Dang S, Chen C, Chua K S. Reinforcement Learning-Driven Dynamic Cache Management with Hotspot-Aware Replacement[J]. 2024, 4(4).

[14] Doshi K, Abudalou S, Yilmaz Y. Tisat: Time series anomaly transformer[J]. arXiv preprint arXiv:2203.05167, 2022.

[15] Athar A, Mozumder M A I, Ali S, et al. Deep learning-based anomaly detection using one-dimensional convolutional neural networks (1D CNN) in machine centers (MCT) and computer numerical control (CNC) machines[J]. PeerJ Computer Science, 2024, 10: e2389.

[16] Zhou T, Ma Z, Wen Q, et al. Fedformer: Frequency enhanced decomposed transformer for long-term series forecasting[C]//International conference on machine learning. PMLR, 2022: 27268-27286.

[17] Zhou Q, Pei C, Sun F, et al. Kan-ad: time series anomaly detection with kolmogorov-Arnold networks[J]. arXiv preprint arXiv:2411.00278, 2024.