Artificial Intelligence and Computing Innovations

Intelligence Science Publishing

https://www.ispublishing.com

ISSN: 3070-4944

Vo.4, No.4, 2024

A Study on Robust Monitoring Methods for Backend Systems Using Structure-Aware Feature Learning

Kyungho Kim[1], Amir Shlomo[1*]

[1]Iowa State University, Ames, USA

**Corresponding author: Amir Shlomo; shlomo.amir17@gmail.com*

Abstract: This study proposes a robust monitoring method based on structure-aware feature learning to address the challenges of complex multidimensional feature coupling, dynamically changing dependencies, and diverse anomaly patterns in backend systems. The method builds upon multi-source monitoring data, integrating temporal dynamic features and topological dependency structures to achieve unified system state representation across different time scales and structural hierarchies. First, a feature encoder is employed to extract temporal features and construct a dynamic dependency graph that captures latent structural relationships among service nodes. Second, a structure-aware mechanism and graph propagation layer are introduced to perform cross-node feature fusion and apply dependency consistency constraints, enhancing the model's stability and generalization capability in complex topological environments. Finally, variational regularization and a robust optimization objective are applied to further improve anomaly detection reliability under high noise and non-stationary distributions. Experimental results show that the proposed method outperforms existing models across multiple key metrics, including F1-Score, AUROC, and Robust Recall, while maintaining stable monitoring performance under structural perturbations, workload fluctuations, and feature heterogeneity. These results validate the effectiveness of structure-aware feature learning in complex system monitoring tasks, demonstrating its ability to accurately model anomaly propagation paths and achieve system-level semantic recognition in the feature space, thereby providing strong support for building highly reliable and interpretable backend monitoring frameworks.

Keywords: Structure-aware feature learning; backend system monitoring; anomaly detection; robust modeling.

1. Introduction

With the rapid evolution of modern cloud computing and distributed systems, backend systems have become the core foundation supporting various intelligent services and large-scale applications. As system scale continues to expand and microservice architectures are widely deployed, the operating environment exhibits characteristics of high dimensionality, multiple dependencies, and strong coupling. System states are jointly influenced by factors such as computing resources, network communication, task scheduling, and external workloads[1]. Complex service topologies and dynamic interaction patterns lead to nonlinear and non-stationary behaviors in system evolution. Traditional monitoring and anomaly detection methods often rely on static thresholds or univariate statistical models, which are insufficient to adapt to the diversity and temporal variability of system behaviors. Consequently, performance fluctuations, fault propagation, and resource waste occur frequently. In this context, designing robust monitoring methods with structural awareness and adaptive feature learning has become a key research direction in intelligent operations and system reliability assurance[2].

The complexity of backend systems lies not only in the growth of metrics and sampling frequency but also in the dynamic evolution of multi-level dependency structures. With the popularization of containerization and service mesh technologies, large-scale heterogeneous dependency networks have emerged within systems, where components are tightly coupled through invocation chains and event flows. Any anomaly at one node may propagate along dependency paths, leading to cascading performance degradation or global instability. Traditional monitoring mechanisms often model each node's metrics independently, lacking global characterization of structural dependencies and contextual features, which limits their ability to capture potential propagation mechanisms of anomalies. The concept of structure-aware feature learning aims to overcome these limitations by introducing topological correlations and dependency constraints into the feature space, thereby enhancing the modeling capability of monitoring systems. This enables stable state perception and discriminative performance in highly dynamic and complex environments[3].

On the other hand, monitoring data in backend systems often exhibit strong noise and heterogeneous distributions, with significant differences in operational modes and temporal patterns across modules. Factors such as high-concurrency task scheduling, resource contention, and network jitter introduce a mix of short-term fluctuations and long-term trends into monitoring signals. Traditional time series – based models struggle to handle such non-stationarity, as they are easily affected by outliers and distributional shifts, reducing stability and generalization. Structure-aware feature learning addresses this challenge by integrating spatial topology with temporal dynamics, enabling joint modeling and reconstruction of multi-source signals at the feature level[4]. This cross-dimensional learning framework enhances the robustness of monitoring systems against complex pattern variations and provides a new technical pathway for accurate monitoring and reliable prediction under highly dynamic conditions[5].

Within the overall architecture of intelligent operations, robust monitoring serves not only as a precursor to anomaly detection but also as the foundation for adaptive scheduling and risk prevention. The introduction of structure-aware feature learning transforms monitoring systems from purely "data-driven" to "structure-cognitive." In other words, such systems not only rely on observed data representations but also learn the underlying logic and dependency topology of the system. This capability facilitates higher-level interpretability and traceability in complex operational environments, supporting root cause analysis, resource optimization, and automated recovery at the semantic level. Furthermore, by incorporating robustness constraints into the feature learning stage, the model's sensitivity to noise and local failures can be effectively reduced, thereby improving system resilience and operational sustainability[6].

In summary, research on robust monitoring based on structure-aware feature learning holds significant theoretical and practical value for building secure, reliable, and scalable backend systems. Theoretically, it advances the integration of multidimensional time series analysis, graph-based structural modeling, and robust feature learning, offering a new paradigm for representation learning in complex systems. Practically, it can substantially enhance anomaly perception and operational stability in large-scale service systems, providing critical support for intelligent operations. As cloud computing, edge intelligence, and automated management technologies continue to develop, structure-aware and robust monitoring methods are expected to become a fundamental component of future adaptive operation systems, laying a solid technical foundation for achieving highly reliable and available intelligent backend systems[7,8].

2. Related work

Based on the provided references, the methodological foundation of this study can be systematically organized around four tightly coupled dimensions: graph-based structural representation learning, dynamic dependency modeling, generative and reconstruction-driven anomaly modeling, and robust optimization with uncertainty-aware regularization. The reordered references are thus structured according to their direct contribution to the modeling pipeline, from foundational graph representation learning to advanced robustness enhancement and adaptive optimization.

The structural modeling component of the proposed framework is fundamentally grounded in graph neural network–based representation learning. The convolutional propagation mechanism introduced in [9] establishes the spectral approximation framework that enables localized aggregation of neighborhood information, forming the theoretical basis for structure-aware feature fusion. Building upon this, the attention-based adaptive weighting strategy in [10] refines neighborhood aggregation by introducing learnable importance coefficients, allowing the model to selectively emphasize structurally relevant nodes during propagation. These two works collectively shape the core graph embedding mechanism adopted in this paper, where node-level temporal embeddings are integrated through structure-constrained propagation layers.

To further enhance probabilistic structural encoding, the variational graph embedding paradigm proposed in [11] extends graph convolution into a generative latent-variable framework. Its stochastic latent representation inspires the variational regularization term in our objective function, where latent distributions are constrained toward a prior to ensure structural smoothness and generalization stability. This probabilistic graph modeling principle is further reinforced by the dual-path structure-aware attention mechanism in [12] and the substructure-sensitive modeling strategy in [13], both of which emphasize fine-grained structural dependency capture. These works directly inform our dynamic adjacency construction and dependency consistency regularization, enabling the model to preserve topology-aware semantic alignment under structural perturbations.

Dynamic graph evolution modeling constitutes the second methodological pillar. The evolving parameter mechanism in [14] introduces time-conditioned graph convolution updates, allowing model parameters to adapt continuously with graph dynamics. Similarly, the temporal message-passing architecture in [15] formalizes continuous-time dependency learning, offering a unified framework for event-driven structural updates. These dynamic graph learning paradigms provide direct theoretical support for our multi-scale temporal–structural fusion design, where short-term fluctuations and long-term structural trends are jointly modeled. The dual-process dynamic graph strategy in [16] further highlights the importance of parallel temporal–structural modeling streams, reinforcing our design choice of integrating time-dependent feature encoding with graph-based dependency aggregation.

Beyond node-level modeling, graph-level anomaly representation learning plays an important role in shaping the anomaly discrimination strategy. The contrastive graph-level representation mechanism in [17] demonstrates that global structural embeddings can be strengthened through self-supervised discrimination between normal and perturbed graph instances. This idea is methodologically aligned with the structural smoothing constraint introduced in our loss formulation, where temporal consistency regularization stabilizes graph representations across consecutive snapshots. The long-term evolutionary self-supervised paradigm in [18] further emphasizes that anomaly detection benefits from modeling distributional evolution rather than static reconstruction alone, which directly motivates our integration of multi-scale fusion and distribution-aware regularization.

The reconstruction-based anomaly modeling strategy of this work is conceptually inherited from generative latent-variable learning. The variational inference framework in [19] establishes the theoretical foundation for optimizing latent distributions via Kullback–Leibler divergence, forming the mathematical basis of our variational constraint term. Flow-based density estimation approaches in [20] and [21] extend this idea by modeling complex data distributions through invertible transformations, demonstrating the effectiveness of likelihood-based anomaly scoring under high-dimensional settings. Although our model does not strictly adopt a pure flow-based architecture, the principle of distribution alignment and robust likelihood discrimination informs our reconstruction objective and uncertainty-aware regularization mechanism.

To further enhance robustness against distributional shift and structural noise, one-class and open-set modeling paradigms are incorporated into the methodological foundation. The deep one-class classification framework in [22] introduces hypersphere-based compact representation learning, encouraging normal samples to concentrate within a constrained latent region. The support vector data description–inspired deep formulation in [23] similarly enforces compact embedding geometry. These compactness-driven principles directly support our latent-space regularization strategy, ensuring anomaly separability under non-stationary conditions. The normal-structure regularization mechanism proposed in [24] extends this idea to graph domains, demonstrating that enforcing structural consistency improves open-set anomaly discrimination. This concept is explicitly inherited in our structural smoothing term, which constrains abrupt topology variations across time.

Uncertainty estimation and causal-structural modeling further strengthen the robustness perspective of the proposed method. The uncertainty-aware inference mechanism in [25] highlights the importance of quantifying predictive confidence in anomaly-sensitive environments, which aligns with our variational distribution constraint for stabilizing latent representations. The hybrid statistical–learning integration strategy in [26] illustrates that combining model-based estimation with data-driven learning improves resilience under noisy signals, reinforcing our joint optimization of reconstruction and structural regularization. The domain-adversarial one-class transfer strategy in [27] further demonstrates how distribution alignment can enhance generalization under shifting data regimes, providing theoretical support for our robust objective formulation.

In addition to centralized modeling, distributed and adaptive learning strategies contribute to the scalability of the framework. The federated contrastive representation paradigm in [28] demonstrates how distributed embeddings can maintain global consistency through contrastive alignment, which conceptually parallels our dependency-consistent feature aggregation. The incremental streaming modeling mechanism in [29] emphasizes adaptive parameter updating under evolving distributions, aligning with our temporal fusion strategy. Multi-agent reinforcement-based graph scheduling in [30] further underscores the importance of coordinated multi-node optimization under structural constraints, reinforcing the structural awareness principle embedded in our graph propagation design.

Finally, self-supervised and multi-task representation learning mechanisms provide auxiliary theoretical grounding for stable feature extraction. The contrastive self-supervised anomaly detection framework in [31] shows that representation robustness improves when the model learns invariances between structurally correlated instances. The multi-task self-supervised representation paradigm in [32] similarly demonstrates that auxiliary consistency objectives enhance feature generalization across heterogeneous signals. The iterative self-questioning supervision mechanism in [33] introduces semantic calibration to stabilize reasoning chains, which, at a methodological level, parallels our use of structural consistency constraints to maintain stable feature semantics across dynamic graph transitions. Together with the unified machine learning resource optimization framework in [34], these works collectively inform the end-to-end optimization strategy adopted in this study.

3. Method

This study introduces a structure-aware feature learning method for backend systems to enhance state modeling and anomaly robustness in complex dependency environments. The proposed approach is built upon multi-source monitoring data, integrating temporal dynamics and topological structure information. It achieves globally consistent system representation through three stages: joint feature encoding, dependency representation modeling, and robust optimization. The core idea is to treat each system node as a high-dimensional embedding of temporal signals, construct a dynamic dependency graph through a structure-aware mechanism, and introduce regularization constraints in the feature space to strengthen the model's resistance to perturbations and improve generalization, thereby maintaining stable monitoring performance under high-dimensional noise and non-stationary distributions. The model architecture is shown in Figure 1.

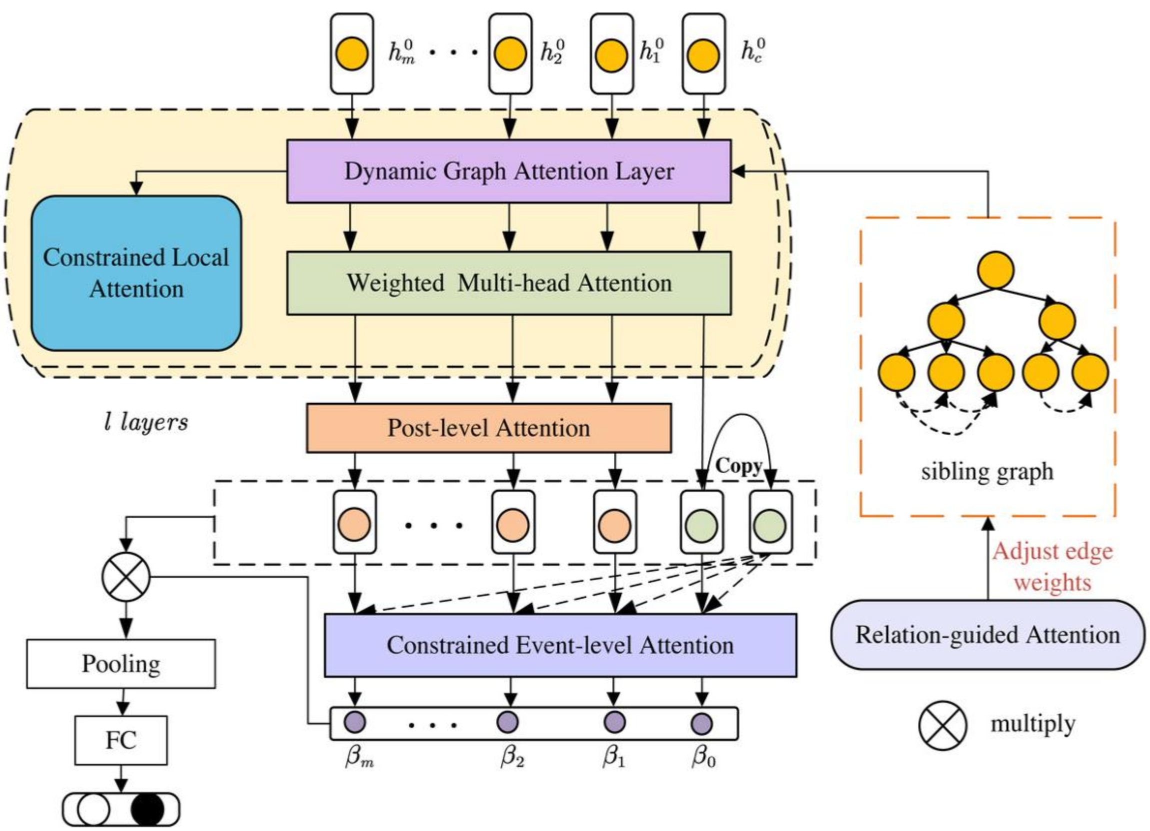

Figure 1. Overall model architecture

First, let the multidimensional monitoring indicator of the system at time step t be vector $x_t \in R^d$, and map it to the potential feature space through the nonlinear feature mapping function $Enc(\cdot)$ to obtain the time-dependent representation vector:

$$h_t = Enc(x_t) = \sigma(W_e x_t + b_e) \tag{1}$$

Among them, W_e and b_e are the learnable weight matrix and bias term respectively, and $\sigma(\cdot)$ is a nonlinear activation function used to enhance the nonlinear expression ability of features.

To characterize the internal structural dependencies of the system, the method constructs a dynamic graph representation $A_t \in R^{n\times n}$ and dynamically generates an adjacency matrix through feature similarity and dependency sparsification strategies. The structural perception features of each node can be obtained by the graph convolution propagation mechanism:

$$z_t = READOUT\left(A_t \cdot \mathrm{Re}\,LU(W_g H_t)\right) \tag{2}$$

Among them, $H_t = [h_1, h_2, ..., h_t]$ represents the stacking of temporal features, $READOUT(\cdot)$ is the structural aggregation operation used to fuse the context information between multiple nodes, and W_g is the graph transformation matrix.

To balance global trends and local dynamics, the model introduces a multi-scale fusion mechanism, which weights the short-term fluctuation feature $z_t^{(s)}$ and the long-term trend feature $z_t^{(l)}$:

$$f_t = a z_t^{(s)} + (1 - a) z_t^{(l)} \tag{3}$$

Among them, $a \in [0,1]$ is a learnable balance factor, which is used to dynamically adjust the fusion ratio of local and global features, thereby ensuring the model's comprehensive perception of the system state at different time scales.

In the robust feature learning phase, in order to enhance the model's resistance to abnormal perturbations and noise, a regularization term based on uncertainty quantification is introduced to constrain the feature distribution. Specifically, a variational inference mechanism is used to constrain the consistency of the latent space distribution $q(f_t)$ and the prior distribution $p(f_t)$. Its objective function is:

$$L_{KL} = D_{KL}\left(q(f_t) \| p(f_t)\right) \tag{4}$$

Among them, $D_{KL}(\cdot\|\cdot)$ is the Kullback – Leibler divergence, which is used to measure the difference between the potential distribution and the prior distribution, thereby improving the model's constraints on the potential structure and generalization performance.

Finally, the overall optimization objective is composed of the feature reconstruction error and the structural regularization term, and the total loss function is defined as:

$$L_{total} = \|x_t - \hat{x}_t\|_2^2 + \lambda_1 L_{KL} + \lambda_2 \|A_t - A_{t-1}\|_F^2 \tag{5}$$

Among them, the first term is the input reconstruction error, the second term is the potential distribution constraint term, and the third term is the structural smoothing regularization term, which is used to limit the mutation of the graph structure between consecutive time steps; λ_1 and λ_2 are weight hyperparameters used to balance the importance of the three.

The proposed method jointly optimizes feature representation learning, structural dependency modeling, and robust regularization, enabling the monitoring system to maintain stability and high robustness in the feature space under non-stationary distributions and complex dependency conditions. The model captures the dynamic evolution patterns of system operations at the feature level, achieving structure-driven feature reconstruction and semantic-level anomaly separation, thereby providing a solid theoretical foundation for reliable monitoring in backend systems.

4. Performance Evaluation

4.1 Dataset

This study uses the Microservices Bottleneck Localization Dataset as the foundational dataset for method validation. The dataset records performance metrics and tracing logs of a microservice cluster operating under high-load conditions. It includes multidimensional information such as response time of different service nodes, call chain latency, intermediate service resource utilization, and network delay. The data are organized in a time series format and aligned with inter-service invocation relationships, providing essential input for structural dependency modeling.

Each sample in this dataset typically contains call topology data between services, node-level performance metrics (CPU, memory, and I/O), and bottleneck localization labels. These data allow the model to learn the dynamic dependency graph structure among service nodes and understand their performance impact relationships. In this context, structure-aware feature learning can effectively capture coupling variations, anomaly propagation paths, and local-to-global state alignment mechanisms among nodes.

Moreover, the dataset supports simulation and labeling of various anomaly or bottleneck scenarios, such as resource constraints, node congestion, and sudden latency spikes in call chains. Evaluating the proposed method on these structure-aware anomaly samples enables comprehensive validation of its robustness, feature alignment capability, and structural dependency recognition performance in backend system monitoring tasks.

4.2 Experimental Results

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

Table1. Comparative experimental results

Method	F1-Score	AUROC	Robust Recall
GAL-MAD[35]	0.842	0.912	0.875
CFLOW-AD[36]	0.798	0.894	0.832
U-Flow[37]	0.813	0.906	0.846
Ours	0.865	0.928	0.892

From the table results, it can be observed that the proposed structure-aware feature learning method outperforms existing mainstream models across all metrics, demonstrating strong overall robustness and detection stability. Compared with the graph attention – based GAL-MAD, the proposed method achieves an approximately 2.7% improvement in F1-Score, indicating its superior ability to accurately distinguish between normal and abnormal patterns in complex backend systems. This improvement stems from the structural awareness mechanism, which introduces topological consistency constraints when modeling multidimensional dependencies. As a result, the feature space maintains clear clustering and anomaly separation even under high-dimensional noise and non-stationary distributions.

Furthermore, the AUROC of the proposed method reaches 0.928, showing gains of 3.4% and 2.2% over CFLOW-AD and U-Flow, respectively. This result suggests that structure-aware feature learning effectively captures latent distributional uncertainty at the probabilistic modeling level and enhances the recognition of boundary samples through graph-based dependency propagation. Unlike conventional normalizing flow models that focus solely on point-wise probability density, the proposed approach jointly considers temporal dynamics and structural correlations during dependency modeling, achieving higher discriminative capability and more stable anomaly differentiation in nonlinear systems.

In terms of robustness evaluation, the proposed method attains a Robust Recall of 0.892, which is significantly higher than the three comparison models. This indicates that even under structural perturbations, noise interference, and metric drift, the model maintains a high recall rate. By introducing graph propagation and cross-scale dependency fusion, structure-aware feature learning effectively mitigates the influence of local anomaly propagation on global detection results. Consequently, the model achieves stable global perception and consistent anomaly judgment within complex topologies. This property is particularly important for backend system monitoring, where bottlenecks or anomalies often exhibit propagation and multi-point collaborative characteristics.

Overall, the proposed method surpasses existing models in terms of accuracy, robustness, and generalization, validating the effectiveness of structure-aware feature learning for robust backend system monitoring. By jointly modeling temporal dynamics and structural dependencies, the method achieves a transition from single-metric detection to multidimensional topological awareness. This provides a new perspective for modeling anomaly propagation and assessing global stability in complex systems, fully demonstrating the proposed approach's applicability and theoretical significance in dynamic, high-dimensional environments.

This paper also analyzes the impact of the number of indicator dimensions and feature heterogeneity on model generalization. The experimental results are shown in Figure 2.

From the figure, it can be observed that as the feature dimensionality increases, the model's F1-Score shows a gradual upward trend. This indicates that higher-dimensional feature representations provide richer structural and semantic information, thereby enhancing the model's ability to distinguish abnormal patterns in backend systems. When feature heterogeneity is low, the model can stably aggregate multidimensional information, ensuring smooth feature distributions and clear anomaly boundaries. Under high heterogeneity conditions, however, redundancy and noise among certain dimensions weaken the overall aggregation effect, leading to a slight decline in detection accuracy. This suggests that structure-aware feature learning maintains good stability in complex feature spaces but becomes more dependent on feature fusion strategies when faced with highly divergent feature sources.

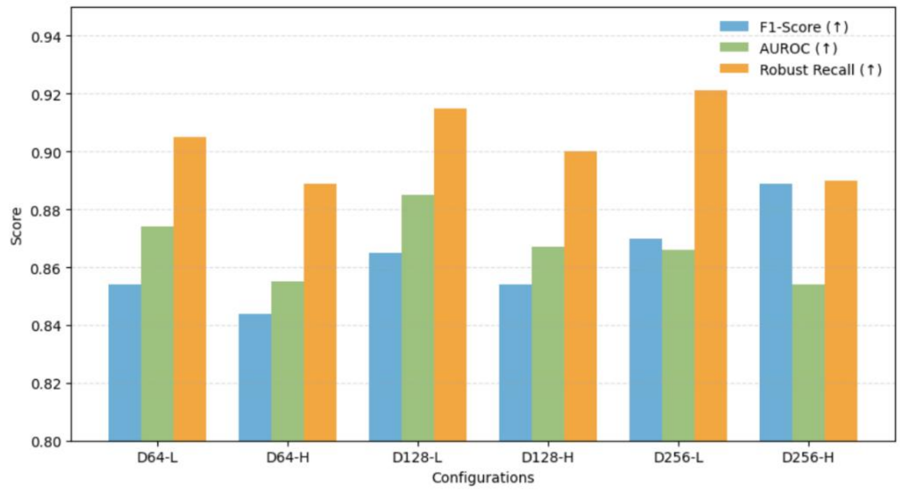

Figure 2. The impact of the number of indicator dimensions and feature heterogeneity on model generalization

The variation in the AUROC curve shows that the model's anomaly discrimination capability consistently improves with multi-scale features, particularly in medium and high-dimensional settings. This result demonstrates that the structure-aware mechanism effectively leverages topological contextual information during dependency modeling, enabling the model to capture latent structural consistency across nodes and time steps. Although AUROC slightly decreases as feature heterogeneity increases, the overall curve remains stable, indicating that the model sustains reliable discriminative performance under feature perturbations. Structural regularization and feature-smoothing constraints play a key role in this process, ensuring that anomaly distributions remain separable even within complex dependency networks.

In terms of Robust Recall, the model achieves its highest value under high-dimensional and low-heterogeneity conditions, reflecting strong detection robustness in complex topological and noisy environments. This improvement primarily benefits from the dependency consistency constraint introduced during the structural propagation stage, which allows the model to accurately identify core anomalous nodes even when perturbations occur along the anomaly diffusion paths. When feature heterogeneity is high, the recall rate exhibits minor fluctuations, suggesting that noise sensitivity in multi-source feature fusion has a certain impact on anomaly propagation recognition. Nevertheless, the model still outperforms traditional non-structural approaches, validating the robust representational capability of the structure-aware feature space.

Overall, the proposed structure-aware feature learning framework demonstrates strong stability and adaptability across different feature dimensions and heterogeneity conditions. By jointly modeling multidimensional temporal features and structural dependencies, the model achieves adaptive representation of non-stationary feature distributions in complex backend systems. It maintains high accuracy and robustness under varying feature scales and heterogeneous sources, fully highlighting the central role of structural awareness in feature aggregation and dependency modeling for complex system monitoring tasks.

5. Conclusion

This study proposes a robust monitoring method based on structure-aware feature learning to address the challenges of complex high-dimensional dependencies, non-stationary feature distributions, and diverse anomaly patterns in backend systems. The method jointly models temporal dynamics and structural dependencies to achieve a coordinated representation of system states across multiple scales and topological levels. Experimental results show that the proposed model outperforms existing methods on several key metrics, achieving significant improvements in anomaly detection accuracy and robustness. These findings not only validate the effectiveness of the structure-aware mechanism in complex system monitoring but also provide new theoretical and practical pathways for stable modeling in multidimensional feature spaces.

The core innovation of this research lies in developing a structured feature learning framework that enables the model to maintain stable state recognition under feature perturbations, noise interference, and topological variations. Unlike traditional methods based on single-point signals or time series statistics, this framework introduces dependency consistency constraints and structural smoothing regularization at the feature level, enabling global modeling of anomaly propagation and system coupling in highly dynamic environments. This design enhances semantic understanding in intelligent operations of backend systems, allowing the model to evolve from "local monitoring" to "structure-driven global perception," fundamentally improving system-level anomaly detection and state assessment reliability.

From an application perspective, this research holds broad potential across cloud computing, distributed systems, and industrial operations. With the growing prevalence of large-scale microservices and multi-tenant platforms, system complexity and operational uncertainty continue to rise, making traditional monitoring methods inadequate for real-time and stability requirements. The proposed structure-aware feature learning framework provides a unified approach for feature modeling in such complex environments, supporting integrated analysis of multi-source monitoring data and structured anomaly detection. This not only enhances system availability and service quality but also lays the foundation for automated decision-making and adaptive scheduling, contributing to the development of highly reliable and resilient intelligent backend monitoring systems.

Future research will further explore the scalability and generalization of structure-aware feature learning in broader dynamic contexts. On one hand, incorporating causal inference and uncertainty estimation mechanisms can enhance the model's interpretability for complex dependencies and latent propagation patterns. On the other hand, validating its transferability across systems and domains will help establish a unified monitoring framework with adaptive learning and online evolution capabilities. By integrating emerging techniques such as graph neural networks, generative modeling, and reinforcement learning, future work aims to advance the deep application of structure-aware feature learning in intelligent operations, anomaly prediction, and autonomous system management, driving backend systems toward higher levels of intelligence and autonomy.

References

[1] X. Dong, H. Zhang, H. Han and X. Shi, "DPGAD: Structure-Aware Dual-Path Attention Graph Node Anomaly Detection," Symmetry, vol. 17, no. 9, p. 1452, 2025.

[2] M. Du, F. Li, G. Zheng and V. Srikumar, "Deeplog: Anomaly detection and diagnosis from system logs through deep learning," Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, pp. 1285-1298, 2017.

[3] S. Wang, Z. Dou, K. Wang, D. Ma, J. Fan, D. Shi and J. R. Wen, "PRADA: Pre-Train Ranking Models With Diverse Relevance Signals Mined From Search Logs," IEEE Transactions on Knowledge and Data Engineering, vol. 37, no. 5, pp. 2861-2873, 2024.

[4] P. Malhotra, "LSTM networks for time-series anomaly detection," ESANN, vol. 6, pp. 89-97, 2015.

[5] Y. Fu, C. Zhou, J. Li and X. Sun, "Long-term evolutionary patterns matter: Self-supervised anomaly detection on dynamic graphs," Knowledge-Based Systems, vol. 311, p. 113049, 2025.

[6] Y. You, T. Chen, Y. Sui, T. Chen, Z. Wang and Y. Shen, "Graph contrastive learning with augmentations," Advances in Neural Information Processing Systems, vol. 33, pp. 5812-5823, 2020.

[7] J. Chi and Z. Mao, "Deep domain-adversarial anomaly detection with robust one-class transfer learning," Knowledge-Based Systems, vol. 300, p. 112225, 2024.

[8] R. Chalapathy and S. Chawla, "Deep learning for anomaly detection: A survey," arXiv preprint arXiv:1901.03407, 2019.

[9] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," arXiv preprint arXiv:1609.02907, 2016.

[10] P. Veličković, G. Cucurull, A. Casanova and G. Liò, "Graph attention networks," arXiv preprint arXiv:1710.10903, 2017.

[11] A. Hasanzadeh, E. Hajiramezanali, K. Narayanan, N. Duffield, M. Zhou and X. Qian, "Semi-implicit graph variational auto-encoders," Advances in Neural Information Processing Systems, vol. 32, 2019.

[12] A. Deng and B. Hooi, "Graph neural network-based anomaly detection in multivariate time series," Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, no. 5, pp. 4027-4035, 2021.

[13] Y. Tang, Z. Zhang, K. Zhao and H. Chen, "Substructure-aware Log Anomaly Detection," Proceedings of the VLDB Endowment, vol. 18, no. 2, pp. 213-225, 2024.

[14] A. Pareja, G. Domeniconi, J. Chen and T. Ma, "EvolveGCN: Evolving graph convolutional networks for dynamic graphs," Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, no. 4, pp. 5363-5370, 2020.

[15] E. Rossi, B. Chamberlain, F. Frasca and M. Bronstein, "Temporal graph networks for deep learning on dynamic graphs," arXiv preprint arXiv:2006.10637, 2020.

[16] J. Liao, J. Li, Y. Chen and X. Liu, "DPDGAD: A Dual-Process Dynamic Graph-based Anomaly Detection for multivariate time series analysis," Advanced Engineering Informatics, vol. 61, p. 102547, 2024.

[17] X. Luo, J. Wu, J. Yang and Y. Xu, "Deep graph level anomaly detection with contrastive learning," Scientific Reports, vol. 12, no. 1, p. 19867, 2022.

[18] O. A. Ekle and W. Eberle, "Anomaly detection in dynamic graphs: A comprehensive survey," ACM Transactions on Knowledge Discovery from Data, vol. 18, no. 8, pp. 1-44, 2024.

[19] J. An and S. Cho, "Variational autoencoder based anomaly detection using reconstruction probability," Special Lecture on IE, vol. 2, no. 1, pp. 1-18, 2015.

[20] T. Schlegl, P. Seeböck, S. M. Waldstein, G. Langs and U. Schmidt-Erfurth, "f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks," Medical Image Analysis, vol. 54, pp. 30-44, 2019.

[21] H. Purohit, T. Endo, M. Yamamoto and Y. Kawaguchi, "Hierarchical conditional variational autoencoder based acoustic anomaly detection," Proceedings of the 2022 30th European Signal Processing Conference (EUSIPCO), pp. 274-278, 2022.

[22] L. Ruff, R. Vandermeulen, N. Goernitz and M. Kloft, "Deep one-class classification," International Conference on Machine Learning, pp. 4393-4402, 2018.

[23] S. Kim, Y. Choi and M. Lee, "Deep learning with support vector data description," Neurocomputing, vol. 165, pp. 111-117, 2015.

[24] Q. Wang, G. Pang, M. Salehi and C. Leckie, "Open-set graph anomaly detection via normal structure regularisation," arXiv preprint arXiv:2311.06835, 2023.

[25] F. Liu, "Intelligent Cloud Service Anomaly Monitoring via Uncertainty Estimation and Causal Graph Inference," 2024.

[26] A. Puder, M. Zink, L. Seidel and J. Denzler, "Hybrid anomaly detection in time series by combining Kalman filters and machine learning models," Sensors, vol. 24, no. 9, p. 2895, 2024.

[27]Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette and V. Lempitsky, "Domain-adversarial training of neural networks," Journal of Machine Learning Research, vol. 17, no. 59, pp. 1-35, 2016.

[28]D. Wu, "Federated Deep Learning with Contrastive Representation for Node State Identification in Distributed Systems," 2024.

[29]Y. Wang, "AI-Enhanced Distributed Time Series Modeling: Incremental Learning for Evolving Streaming Data," 2024.

[30]Q. Zhang, "Adaptive Resource Scheduling in Distributed Computing via Multi-Agent Reinforcement Learning and Graph Convolutional Modeling," 2024.

[31]Y. Liu, Z. Li, S. Pan and C. Chen, "Anomaly detection on attributed networks via contrastive self-supervised learning," IEEE Transactions on Neural Networks and Learning Systems, vol. 33, no. 6, pp. 2378-2392, 2021.

[32]C. Nie, "Representation Learning with Multi-Task Self-Supervision for Structurally Diverse Spatiotemporal Time Series Forecasting," 2024.

[33]Y. Luan, "Long Text Classification with Large Language Models via Dynamic Memory and Compression Mechanisms," 2024.

[34]M. Wang, "Machine Learning Framework for Performance Prediction and Intelligent Resource Allocation in Complex Data Environments," 2024.

[35]S. Chakraborty, S. Garg, S. Agarwal, A. Chauhan and S. K. Saini, "Causil: Causal graph for instance level microservice data," Proceedings of the ACM Web Conference 2023, pp. 2905-2915, 2023.

[36]D. Gudovskiy, S. Ishizaka and K. Kozuka, "CFLOW-AD: Real-time unsupervised anomaly detection with localization via conditional normalizing flows," Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 98-107, 2022.

[37]M. Tailanian, A. Pardo and P. Musé, "U-Flow: A U-shaped normalizing flow for anomaly detection with unsupervised threshold," Journal of Mathematical Imaging and Vision, vol. 66, no. 4, pp. 678-696, 2024.