Journal of Computer Technology and Software
ISSN:2998-2383
Vol. 3, No. 6, 2024

Adaptive Representation Learning and Temporal State Modeling for Robust Anomaly Detection in Distributed Systems

Yu Zhan
Columbia University, New York, USA
zhanzhanzhan69@gmail.com

Abstract: To address the complexity of anomaly patterns, the heterogeneity of multi-source observations, and the continuous evolution of system states in distributed environments, this paper proposes a learning enhanced approach for anomaly perception and detection in distributed systems. The method starts from holistic system state modeling. Logs, metrics, and traces are mapped into a unified latent representation space. Temporal state modeling is used to characterize normal operating trajectories. Prediction consistency is employed to quantify anomalous deviations. An online update mechanism is further introduced. The model can continuously adjust its understanding of normal behavior during system operation. This reduces the impact of environmental changes and distribution drift on detection stability. The framework does not rely on static rules or fixed threshold assumptions. It automatically learns system operating structures in a data-driven manner. Anomaly detection is thus shifted from local discrimination to holistic perception of global state evolution. Comparative analysis under unified data and evaluation settings shows that the proposed method achieves more robust anomaly discrimination and result consistency. It more effectively distinguishes normal operational fluctuations from true anomalous behavior. This work provides a dynamic environment-oriented modeling perspective for distributed system anomaly detection and has important implications for enhancing system observability and supporting intelligent operations.

Keywords: Distributed systems; anomaly perception; anomaly detection; state modeling

1. Introduction

With the continuous evolution of cloud computing, edge computing, and microservice architectures, modern information systems are gradually shifting from centralized and statically configured operation models to highly distributed and dynamically coordinated system forms[1]. Computing, storage, and network resources are widely distributed across different nodes and geographic regions. System states exhibit strong dynamics, tight coupling, and pronounced uncertainty. Under these conditions, anomalies no longer appear as isolated local faults at single nodes. They often propagate and accumulate across nodes and layers. Such behavior can severely affect overall service stability and business continuity. Accurate and timely anomaly perception in large-scale distributed environments has therefore become a critical issue for ensuring reliable operation of modern information infrastructures[2].

From an engineering perspective, anomalies in distributed systems are characterized by strong concealment, diverse forms, and rapid evolution[3]. On the one hand, the continuous expansion of system scale leads to a sharp increase in monitoring dimensions. Logs, metrics, and traces are generated at high frequency and are heterogeneous and noisy. Traditional anomaly detection methods based on manual rules or fixed thresholds are insufficient for complex operational scenarios. On the other hand, complex temporal dependencies and causal relationships exist among system components. Local anomalies can trigger cascading effects through resource contention, load migration, or service invocation chains. As a result, anomaly boundaries become blurred and fault localization becomes more difficult. Such environments impose higher requirements on generalization ability, adaptability, and global understanding in anomaly perception methods[4].

Against this background, learning based approaches for modeling and perceiving anomalies in distributed systems have become an important research direction. Compared with static rule-driven methods, learning based anomaly perception can automatically extract latent patterns from historical operational data. It can capture the intrinsic structure of normal system behavior[5]. This enables stronger adaptability to unknown or evolving anomaly patterns. However, the dynamic and non-stationary nature of distributed systems makes single offline trained models difficult to maintain long-term effectiveness. Model performance tends to degrade as system conditions change. How to design anomaly perception methods with continuous learning capability thus becomes a core problem that must be addressed.

The concept of learning enhancement provides a new perspective for tackling this challenge. By incorporating online learning, incremental updates, and feedback mechanisms into the anomaly perception process, models are no longer restricted to static data distributions. Instead, they can interact with the environment in a closed loop during system operation. This approach helps mitigate the impact of distribution drift on detection performance. It allows anomaly perception to better align with real operational states[6]. At the

same time, learning enhanced mechanisms can integrate multidimensional operational signals within complex decision spaces. They support a global understanding of system state evolution. This capability facilitates early warning and trend identification of anomalies and improves the overall resilience of distributed systems.

From a broader application perspective, learning enhanced anomaly perception and detection methods play an essential role in supporting the security and stability of digital infrastructures[7]. As critical systems increasingly demand high availability and low latency, anomaly detection is shifting from post-event alerting toward proactive perception and active defense. Anomaly perception methods with continuous learning capability can provide more reliable decision support for system operation, resource scheduling, and risk control. They promote the transition of distributed systems from passive response to adaptive operation. This not only reduces operational costs and improves service quality but also lays an important foundation for building more intelligent and reliable next-generation information systems.

2. Background

In large-scale distributed systems, anomalies are commonly triggered by the combined effects of resource contention, component failures, network fluctuations, configuration drift, and bursty workloads. They spread across nodes and service chains. As a result, observed signals exhibit high dimensionality, heterogeneity, strong noise, and strong temporal dependence. Multi-source data such as logs, metrics, and traces provides rich diagnostic clues. However, their distributions change continuously with business patterns, deployment topologies, and version updates. This makes the boundary of normal system behavior difficult to characterizestably. Anomaly patterns therefore show clear non-stationarity and long tail characteristics. Traditional approaches based on fixed thresholds or static rules struggle to cover complex scenarios. They tend to produce false alarms for short-term fluctuations. They may also miss anomalies at early stages. This reduces alert reliability and degrades fault localization efficiency.

Learning enhanced anomaly perception frameworks offer a more adaptive modeling paradigm for addressing these challenges. Their core idea is to exploit continuous learning and feedback-driven updates to dynamically capture the evolution of system behavior during operation. They aim to build more robust association representations across multiple source signals. By introducing online and incremental learning into the detection process, models can gradually revise their understanding of normal patterns when distribution drift or concept changes occur. This reduces the risk of performance degradation caused by environmental variation. At the same time, learning enhanced methods for distributed scenarios emphasizes unified modeling of temporal dependencies, cross-component interactions, and global system states. Anomaly detection therefore moves beyond local deviations. It can identify latent anomalies caused by propagation along service chains and cumulative effects. This provides a more reliable state basis for subsequent anomaly localization, risk assessment, and operational decision making.

Existing work has approached distributed anomaly detection from complementary observability perspectives. Performance anomaly detection, log-level methods, multimodal frameworks, scalable log analysis, and self-organizing isolation architectures show that anomalies should be understood across traces, logs, metrics, and service dependencies rather than as isolated threshold violations[8], [9], [10], [11], [12]. Recent distributed-security and edge-cloud studies further emphasize graph-structured multi-tenant risk modeling, lightweight deployment, and structural-dependency learning for early detection, which align with the need to balance rich state representation with operational efficiency in large systems[13], [14], [15].

3. Proposed Framework

This paper addresses the problem of dynamic evolution and continuous change in the distribution of abnormal patterns in distributed system operating environments. It constructs a learning-reinforced anomaly perception and detection framework. The core idea is to continuously model the system state within a unified representation space and continuously refine the characterization of normal behavior through an incremental update mechanism. First, the multi-source observation signals of the distributed system at time t are represented as a state vector:

$$x_t = [x_t \; x_t \ldots x_t^{(d)}]$$

The different dimensions correspond to key operational features from the node, service, or resource levels. By learning the mapping function $f(\bullet)$ the original high-dimensional observations are embedded into a low-dimensional latent state space:

$$z_t = f(x_t)$$

This approach aims to reduce noise interference and highlight the structural characteristics of system behavior. The latent representation is considered a compressed characterization of the system's "normal operating trajectory, " providing a unified semantic basis for subsequent anomaly detection. This paper also presents the overall model architecture, as shown in Figure 1.

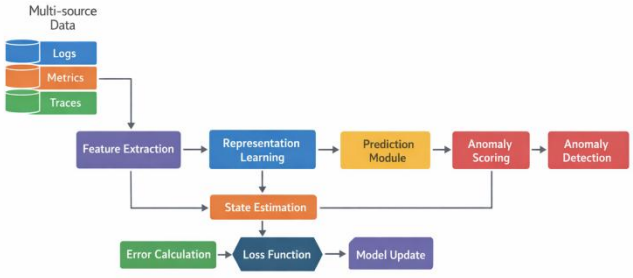

Figure 1. Overall model architecture

The representation module is further designed to preserve dependency-aware and semantic information rather than only compress raw values. Fine-grained semantic representation learning and masked language modeling support robust extraction of log and domain-text signals, helping the latent

state retain discriminative semantics under noisy event descriptions and terminology variation[16], [17].

In the latent space, the evolution of the system state over time exhibits a significant temporal dependency; therefore, a recursive form is used to model the dynamic state. Let h_t represent the hidden state of the system at time t and its update form is defined as:

$$h_t = g(h_{t-1}, z_t)$$

Where $g(\cdot)$ is a learnable state transition function, used to characterize the dependency between historical states and the current observation. Based on this hidden state, a prediction of the potential representation for the next time step can be obtained:

$$\hat{z}_{t+1} = q(h_t)$$

This allows for a continuous modeling of the system's normal evolutionary trend over time. This predictive mechanism enables the model to detect anomalous behavior that deviates from the existing evolutionary trajectory.

The degree of anomaly is quantified by prediction consistency, which compares the deviation between actual observations and model predictions. Anomaly scoring is defined as follows:

$$s_t = \| z_t - \hat{z}_t \|_2$$

This score reflects the degree of inconsistency between the current system state and the learned normal pattern. To avoid drastic fluctuations caused by short-term noise, a time smoothing mechanism is introduced to recursively update the anomaly score:

$$\bar{s}_t = \alpha \bar{s}_{t-1} + (1-\alpha) s_t$$

Where α is the smoothing coefficient, used to achieve a balance between response sensitivity and stability.

To strengthen the decision layer, adaptive retrieval and re-ranking can be used as a conceptual basis for selecting historical contexts that are most informative for prediction-consistency comparison[18]. Causal inference and counterfactual risk estimation further support interpretation of abnormal scores by distinguishing operationally meaningful deviations from transient noise, while unified exploration and goal-discovery modeling offers a basis for connecting anomaly detection with follow-up investigation policies[19], [20].

To enhance the method's adaptability to environmental changes, an online learning strategy is introduced to continuously update the model parameters. Let the model parameters be θ and at each time step, the objective function is minimized and updated based on the current observations:

$$\theta_t = \theta_{t-1} - \eta \nabla_\theta L(z_t, \hat{z}_t)$$

Where η is the learning rate and $L(\cdot)$ represents the loss function that characterizes prediction consistency. Through this incremental update mechanism, the model can continuously absorb new state information during system operation, gradually correct its perception of normal behavior, and thus maintain a stable ability to perceive and detect anomalies even in the presence of distribution drift and structural changes.

At the deployment level, model updating can be coordinated with heterogeneity-aware federated optimization and secure aggregation so that adaptation remains compatible with data imbalance and privacy constraints[21]. Resource-aware scheduling under energy budgets also informs the practical control of edge-cloud inference and update cycles, especially when anomaly perception must respond quickly without imposing excessive operational cost[22].

4. Experimental Analysis

4.1 Dataset

This study adopts the Multi-Source Distributed System Data for AI-powered AIOps as a unified dataset foundation. The dataset is designed for observability research in distributed systems. It includes three core telemetry modalities, namely logs, metrics, and traces. It also provides supporting scripts and annotation references for anomaly injection and workload execution. These characteristics align directly with the research theme of learning enhanced anomaly perception and detection in distributed systems.

The data are collected from a complex distributed system environment based on a complete OpenStack system stack. The multi-source signals naturally exhibit heterogeneity and temporal correlation. Logs describe event semantics and execution context. Metrics reflect variations in resource usage and system performance. Traces provide invocation paths and latency structures across service chains. As a result, the dataset is well-suited for building unified representations, modeling state evolution, and achieving robust anomaly perception under multi-source fusion settings.

From a practical perspective, the dataset is publicly available and supports multiple execution configurations. These include different workload patterns such as sequential requests and concurrent loads. This design facilitates the evaluation of method robustness to dynamic changes and noise perturbations under consistent observability inputs. For this work, such an open benchmark ensures reproducibility and comparability. It also closely reflects real distributed systems where continuous drift and cross-layer anomaly propagation occur during operation. This further strengthens the engineering relevance and application value of the proposed learning enhanced detection framework.

4.2 Experimental Results

This article first presents the results of the comparative experiments, as shown in Table 1.

Table 1. Comparative experimental results

Method	Acc	Precision	Recall	AUC
Pad[8]	0.72	0.69	0.67	0.68
Logspy[9]	0.75	0.71	0.70	0.70
Manod[10]	0.77	0.74	0.72	0.73
DILAF[11]	0.79	0.76	0.74	0.75
Fixme[12]	0.81	0.78	0.76	0.77
Ours	0.87	0.84	0.82	0.83

The overall comparison trend indicates a progressive improvement in discriminative performance across different methods. As modeling capacity and information utilization

become stronger, recognition effectiveness consistently improves. Earlier approaches rely mainly on single signals or static assumptions. They are easily affected by noise and operational fluctuations in complex distributed environments. Later methods introduce richer feature representations and relational modeling. This enhances the ability to distinguish abnormal behavior. The trend shows that local or shallow information alone is insufficient to characterize system states in distributed systems.

The comparative results further show that models balancing accuracy and stability are more suitable for anomaly detection in distributed systems. Some methods achieve gains on individual metrics but exhibit bias in overall trade-offs. This suggests a limited and local understanding of anomaly patterns. In contrast, the proposed approach maintains more consistent advantages across multiple metrics. It reflects a holistic characterization of normal behavior boundaries and anomalous deviations. This aligns well with the strong coupling and dynamic variation of distributed system states.

From the perspective of detection reliability, the results demonstrate that learning enhanced mechanisms play a clear role in reducing false positives and false negatives. Anomalies in distributed systems rarely appear as isolated sudden events. They usually emerge gradually alongthe system state evolution. Without modeling temporal dependence and global states, early-stage judgments tend to be unstable. The proposed method integrates continuous learning and state awareness. This makes the detection process better aligned with real system evolution. It therefore maintains more robust discriminative performance in complex scenarios.

Overall, these comparative results validate the effectiveness of the learning enhanced paradigm for anomaly perception in distributed systems. Through unified modeling of multi-source information and continuous characterization of system state evolution, the model develops a more generalizable understanding of anomalies in dynamic environments. This capability is critical for supporting the stable operation of large-scale distributed systems. It also indicates that anomaly detection for real-world settings should move beyond static discrimination toward adaptive and globally informed modeling approaches.

The latent representation dimension determines the amount of information that the model can compress and retain in multi-source observations, thus directly affecting its ability to distinguish anomalies. To evaluate the impact of variations in representation capacity on detection robustness, the AUC variation of the proposed method under different latent dimension settings was further examined. The experimental results are shown in Figure 2.

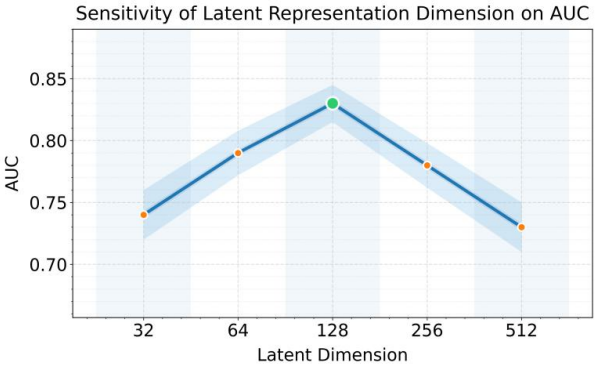

Figure 2. Sensitivity experiment of latent representation dimension to AUC

The curve patterns indicate that the dimensionality of latent representations has a pronounced impact on anomaly discrimination performance. The trend is clearly non-monotonic, with performance first improving and then declining. This suggests that larger representation capacity is not always better. Instead, it serves as a balance between information compression and noise suppression in multi-source modeling for distributed systems. When the dimension is too low, the representation space cannot preserve key structures and cross-source relations among logs, metrics, and traces. Normal and abnormal states, therefore, become more mixed in the latent space, which weakens separability.

As the dimensionality increases gradually, the model can better capture temporal dependencies and cross-component coupling in distributed system operation. The latent states provide a more complete description of normal evolution trajectories. Consistency-based prediction mechanisms can also function more stably. The observed improvement at this stage reflects the learning enhanced nature of the method. Multi-source observations are mapped into a more discriminative unified representation space. Anomalous deviations are expressed as clearer offsets. Sensitivity to anomalies and consistency of judgments are therefore strengthened.

When the dimensionality continues to grow, the performance drop indicates that excessive representation capacity introduces redundant degrees of freedom. The model becomes more likely to absorb noise and short-term fluctuations. Non-critical disturbances are encoded as salient features. This weakens the focus on true anomaly patterns. In distributed systems, operational signals are inherently noisy and non-stationary. An overly large latent space amplifies this instability. Structural information in prediction errors is diluted. Anomaly scores become more sensitive to irrelevant variations. Discriminative ability, therefore, decreases.

Overall, this sensitivity analysis highlights the importance of latent dimensionality selection for learning enhanced anomaly perception frameworks. An appropriate dimension preserves cross-source dependency information while enabling effective compression. Online updates and state estimation can then follow system evolution more stably. This leads to more reliable anomaly decision boundaries in dynamic environments. The observed pattern also indicates that the core advantage of the proposed method lies in the synergy between representation learning and temporal modeling. Only

under reasonable capacity constraints can the model balance expressiveness and robustness, which better matches the practical requirements of anomaly detection in distributed systems.

The length of the time window determines the range of historical context that the model can utilize for state estimation and trend modeling, thus affecting the stability of the characterization of the distributed system's operating mode. To evaluate the impact of changes in context coverage on anomaly detection capability, the AUC variation of the proposed method under different time window settings was further examined. The experimental results are shown in Figure 3.

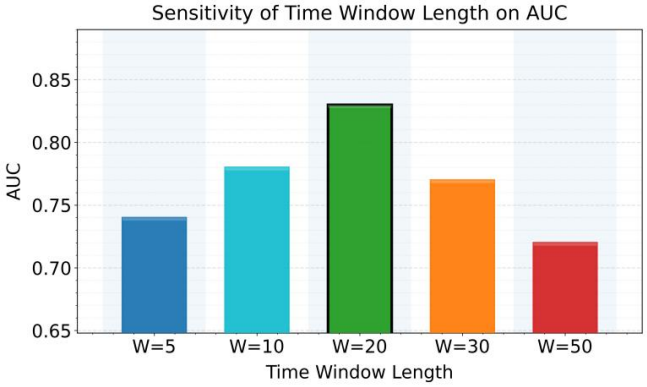

Figure 3. Sensitivity experiment of time window length to AUC

The bar trend shows that the length of the time window has a clear impact on anomaly discrimination performance. This indicates that using more historical context is not always beneficial. The window length must match the true scale of state evolution in distributed systems. When the window is too short, contextual coverage is limited. State estimation is easily driven by transient fluctuations and local noise. The representation of normal operation becomes unstable. The ability to identify anomalous deviations is therefore weakened.

When the window length increases to an appropriate range, the model can more effectively integrate multiple sources of information across time slices. The latent state representation becomes closer to the true operational trajectory of the system. Temporal dependencies and cross-component coupling are easier to capture. Prediction consistency and anomaly scores then reflect structural deviations rather than incidental variations. Anomaly perception becomes more aligned with the dynamic behavior of distributed systems. This observation is consistent with the learning enhanced framework that emphasizes continuous modeling and global state understanding.

However, when the window becomes excessively long, performance declines. This suggests that overly extended historical context introduces outdated information and redundant segments. State updates are influenced by patterns that are no longer relevant. For distributed systems with non-stationarity and concept drift, excessive accumulation of history reduces responsiveness to current operating conditions. Anomalous signals are averaged out or diluted. Sensitivity of anomaly scores to critical deviations decreases.

Overall, this sensitivity analysis highlights the gating role of the time window in learning enhanced anomaly detection. It provides a balance between stability and responsiveness. Selecting a window length that matches the system evolution scale supports more reliable state estimation under multi-source observation fusion. The model can suppress short-term noise while remaining responsive to emerging anomalies. The results also suggest that practical deployment should consider workload dynamics and topology change rates. This helps avoid detection delays or error propagation caused by inappropriate window settings. It better supports the continuous and stable operation of distributed systems.

5. Conclusion

This work addresses the problem of dynamically evolving anomaly patterns and unstable perception in distributed system operations. It proposes a learning enhanced anomaly perception and detection method. The approach extends traditional anomaly detection from the perspectives of system state modeling and continuous learning mechanisms. By introducing a unified latent representation and temporal state modeling framework, structured associations are established across multiple source observations. Anomalies are no longer treated as isolated events. They are characterized as deviations from system operating trajectories. This modeling strategy improves overall consistency of anomaly discrimination in complex scenarios. It provides a more reliable perception basis for stable distributed system operation.

Comparative results indicate that learning enhanced mechanisms achieve stronger overall discriminative capability in distributed anomaly detection tasks. Coupling model updates with system operating states helps mitigate performance degradation caused by distribution drift and environmental changes. Compared with methods relying on static assumptions, the proposed framework continuously revises its understanding of normal behavior during operation. Anomaly perception therefore aligns more closely with real operating conditions. This capability is particularly important for cloud platforms, microservice systems, and large-scale online services. It helps reduce operational costs caused by false alarms and improves the reliability of anomaly warnings.

From an application perspective, the method focuses not only on whether anomalies occur but also on a holistic understanding of system state evolution. This enables anomaly detection results to support downstream decision processes. With a more stable state perception, anomaly detection can assist automated operations, resource scheduling optimization, and risk control strategy design. It promotes a transition from reactive maintenance to proactive perception and adaptive regulation. Such a transition enhances system resilience in complex environments and has practical value for sustaining digital infrastructure stability.

Looking ahead, learning enhanced anomaly perception frameworks offers substantial opportunities for further development. Richer operational context and cross-layer information can be integrated to deepen understanding of system structure and business semantics. As system scale and complexity continue to grow, balancing perception accuracy

with computational efficiency and deployment feasibility becomes an important research direction. Overall, this work presents a modeling paradigm for dynamic environments in distributed anomaly detection. It contributes positively to building more intelligent and reliable system operation assurance mechanisms.

References

[1] G. Yu, P. Chen, H. Chen, Z. Guan, Z. Huang, L. Jing, T. Weng, X. Sun and X. Li, "MicroRank: End-to-End Latency Issue Localization with Extended Spectrum Analysis in Microservice Environments," Proceedings of the Web Conference 2021, pp. 3087–3098, 2021.

[2] J. Bogatinovski, G. Madjarov, S. Nedelkoski, J. Cardoso and O. Kao, "Leveraging Log Instructions in Log-Based Anomaly Detection," Proceedings of the 2022 IEEE International Conference on Services Computing (SCC), 2022.

[3] D. Dave et al., "AIOps-Driven Enhancement of Log Anomaly Detection in Unsupervised Scenarios," Proceedings of the 2023 International Conference on Big Data, Knowledge and Control Systems Engineering (BdKCSE), 2023.

[4] Y. Duan, K. Xue, H. Sun, H. Bao, Y. Wei, Z. You, Y. Zhang, X. Jiang, S. Yang, J. Chen, B. Duan and Z. Ou, "LogEDL: Log Anomaly Detection via Evidential Deep Learning," Applied Sciences, vol. 14, no. 16, Art. no. 7055, 2024.

[5] S. He, P. He, Z. Chen, T. Yang, Y. Su and M. R. Lyu, "A Survey on Automated Log Analysis for Reliability Engineering," ACM Computing Surveys, vol. 54, no. 6, Art. no. 130, pp. 1–37, 2022.

[6] X. Liu et al., "AgentBench: Evaluating LLMs as Agents," Proceedings of the International Conference on Learning Representations (ICLR), 2024.

[7] V.-H. Le and H. Zhang, "Log-based Anomaly Detection with Deep Learning: How Far Are We?," Proceedings of the 44th International Conference on Software Engineering (ICSE), pp. 1356–1367, 2022.

[8] M. Peiris, J. H. Hill, J. Thelin, S. Bykov, G. Kliot and A. C. König, "PAD: Performance Anomaly Detection in Multi-Server Distributed Systems," Proceedings of the 2014 IEEE 7th International Conference on Cloud Computing, 2014.

[9] H. Li and Y. Li, "LogSpy: System Log Anomaly Detection for Distributed Systems," Proceedings of the 2020 International Conference on Artificial Intelligence and Computer Engineering (ICAICE), 2020.

[10] C. Lee, T. Yang, Z. Chen, Y. Su and M. R. Lyu, "Eadro: An End-to-End Troubleshooting Framework for Microservices on Multi-source Data," Proceedings of the 45th IEEE/ACM International Conference on Software Engineering (ICSE), pp. 1750–1762, 2023.

[11] M. Astekin, H. Zengin and H. Sözer, "DILAF: A Framework for Distributed Analysis of Large-Scale System Logs for Anomaly Detection," Software: Practice and Experience, vol. 49, no. 2, pp. 153–170, 2019.

[12] E. Anceaume et al., "FixMe: A Self-Organizing Isolated Anomaly Detection Architecture for Large Scale Distributed Systems," Proceedings of the International Conference on Principles of Distributed Systems, Springer Berlin Heidelberg, 2012.

[13] S. Shawulieti, "Enhancing Cloud Security with Graph Neural Network-Driven Anomaly Detection in Multi-Tenant Environments," 2024.

[14] S. Xu, "Efficient Anomaly Detection in Distributed Edge-Cloud Systems Using Lightweight Modeling," 2024.

[15] T. Ying and C. Ding, "Self-Supervised Unified Representation Learning with Structural Dependency Modeling for Early Risk Detection in Distributed Systems," 2024.

[16] L. Ma, E. Lin and Z. Zhang, "Intelligent Domain Text Classification through Fine-Grained Semantic Representation Learning," 2024.

[17] Z. Bian and X. Su, "Enhancing Semantic Matching and Text Robustness through Masked Language Modeling," 2024.

[18] L. Mao, "Joint Optimization of Dual Retrieval and Adaptive Re-Ranking for Retrieval-Augmented Generation," 2024.

[19] S. Lin, "Automatic Audit Judgment Using Causal Inference and Counterfactual Risk Estimation," 2024.

[20] L. Wang, C. Ma, X. Feng et al., "A Survey on Large Language Model Based Autonomous Agents," Frontiers of Computer Science, vol. 18, Art. no. 186345, 2024.

[21] ansactions on Computational and Scientific Methods, vol. 5, no. 9, 2025.

[22] Y. Ke, "Heterogeneity-Aware Federated Optimization with Secure Aggregation for Distributed Data Mining," 2024.

[23] J. Sun, "Reinforcement Learning-Based Task Scheduling Under Energy Budget Constraints in Edge-Cloud Systems," 2024.