Artificial Intelligence and Computing Innovations

Intelligence Science Publishing

https://www.ispublishing.com

ISSN: 3070-4944

Vo.1, No.2, 2021

Contrastive Representation Learning for Anomaly Detection in Cloud-Based Backend Services

Bruce Barlocker[1], Xu Yan[2*]

[1]Cornell University

[2]Independent Researcher

**Corresponding author: Xu Yan; yyybbb1998@gmail.com*

Abstract: This study proposes an anomaly detection method based on contrastive representation learning to address the challenges of complex multi-metric coupling, dynamic dependency variation, and label scarcity in backend service systems. The method first employs a multi-scale temporal encoding module to extract dynamic features at different time granularities, capturing both short-term fluctuations and long-term trends in system operation. An adaptive dependency modeling mechanism is then constructed, which generates dynamic graph structures through feature similarity projection to characterize semantic relationships and temporal dependencies among services. On this basis, a contrastive learning objective is introduced, where structured constraints between positive and negative sample pairs enable stable feature aggregation and boundary separation under unsupervised or weakly supervised conditions. A temporal consistency regularization term is also incorporated to maintain state smoothness across time steps, enhancing robustness under non-stationary distributions. The proposed method is validated on multidimensional cloud workload datasets, and results show that it achieves high detection accuracy and stability even with incomplete labels and noisy pseudo-labels. Overall, the method not only realizes self-supervised and structure-aware anomaly detection but also demonstrates strong performance in feature representation, anomaly discrimination, and distribution shift adaptation, providing an efficient and general solution for intelligent backend operations and multi-source time-series analysis.

Keywords: Contrastive representation learning; anomaly detection; multi-scale time series modeling; cloud-based backend systems.

1. Introduction

This study focuses on a contrastive representation learning approach for anomaly detection in backend services. The research background stems from the rapid proliferation of cloud-native and microservice architectures. As application scale and complexity continue to grow, backend systems have become highly dynamic, strongly dependent, and multidimensionally coupled. Thousands of service instances and components collaborate asynchronously, forming a vast operational network. During system operation, performance fluctuations, request delays, resource contention, and call chain anomalies frequently occur, posing serious threats to overall service stability. Traditional detection methods based on thresholds, statistical features, or univariate time series analysis struggle to adapt to such high-dimensional dynamic environments. They fail to capture cross-component, cross-temporal, and cross-context anomaly patterns accurately, leading to detection delays, high false alarm rates, and difficulties in localization. Therefore, exploring detection methods with stronger generalization and structural understanding capabilities has become a key issue in intelligent backend operations[1].

In modern backend systems, metric data exhibit strong diversity and non-stationarity. Multiple monitoring signals, such as CPU utilization, memory usage, request latency, I/O throughput, and container load, together form a complex temporal network. These signals contain both short-term high-frequency fluctuations and long-term low-frequency structures. Meanwhile, service dependencies, load balancing strategies, and environmental noise continuously alter statistical correlations among metrics, making anomaly patterns heterogeneous and dynamically transferable across scenarios[2]. In this context, extracting semantically stable representations from massive and noisy data becomes a critical challenge for accurate anomaly detection. Traditional supervised models rely on large-scale labeled datasets. However, in real-world systems, anomalies are rare and often subject to distribution shifts, limiting the feasibility of supervised learning. Therefore, designing a representation learning mechanism that requires few labels and can adaptively capture latent semantic dependencies across services is crucial for efficient anomaly detection[3].

Contrastive representation learning, as a self-supervised feature modeling paradigm, offers a new perspective for anomaly detection in backend systems. Its core idea is to construct contrastive tasks to explore intrinsic similarities and differences within data under label-free conditions. This enables the model to automatically learn discriminative latent features for normal patterns. In backend environments, monitoring data from different time windows, nodes, or contexts can be aligned through contrastive constraints, producing semantically consistent representations[4]. This helps the model maintain representational stability under non-stationary distributions and environmental perturbations. Compared with reconstruction-based or prediction-based self-supervised approaches, contrastive learning emphasizes structural consistency and global separability, effectively reducing redundancy in high-dimensional temporal features. This self-adaptive contrastive mechanism grounded in feature distribution structures establishes a theoretical foundation for unsupervised anomaly identification in complex dynamic systems[5].

In practical scenarios of intelligent operations and reliability management, the introduction of contrastive representation learning not only changes the way features are extracted but also transforms the detection paradigm. Traditional anomaly detection often relies on threshold-based judgments or clustering-based outlier scoring of single-point signals, which cannot handle collaborative anomalies caused by dynamic service dependencies. By contrast, contrastive learning explicitly models the representation distribution of normal operation patterns in feature space, enabling the system to identify states that deviate from this distribution. Furthermore, contrastive mechanisms can be integrated with graph-based modeling and attention structures to enhance feature discrimination from both topological and temporal perspectives. This hybrid approach improves adaptability to complex dependency structures and provides a semantic foundation for higher-level operational tasks such as root cause analysis and service impact prediction[6].

Overall, research on contrastive representation learning for backend service anomaly detection holds significant theoretical and practical value. Theoretically, it extends self-supervised learning into the domain of complex system time-series modeling, promoting structural understanding of non-stationary, high-dimensional, and interdependent feature spaces. Practically, it provides intelligent support for achieving high reliability in cloud-native and distributed systems. By achieving semantically consistent representation and efficient anomaly identification of system behaviors, this approach enhances backend service stability, reduces maintenance costs, and lays the foundation for future automated and self-healing operation systems. Therefore, developing contrastive representation learning-based anomaly detection methods for backend services is not only a natural trend in technological evolution but also an essential step toward sustainable intelligent service ecosystems[7].

2. Method

This study introduces a method called Contrastive Representation Learning for Backend Service Anomaly Detection (CRL-BSAD), which aims to achieve semantic-consistent modeling and anomaly representation of multidimensional system metrics under unsupervised conditions. The core idea of this approach is to employ a contrastive learning mechanism to measure the representational discrepancy between "normal" and "abnormal" sample pairs within a multi-source temporal feature space, thereby identifying potential anomaly patterns at the feature distribution level. The overall model consists of three main modules: a multi-scale feature encoding layer, a semantic consistency modeling layer, and a contrastive constraint optimization layer. The feature encoding layer extracts both local dynamics and global trends from multidimensional metric sequences; the semantic modeling layer aggregates cross-service features through context-dependent graph construction and attention mechanisms; and the contrastive optimization layer minimizes representational inconsistency via a structured constraint function to produce robust anomaly-aware embeddings. The model architecture is shown in Figure 1.

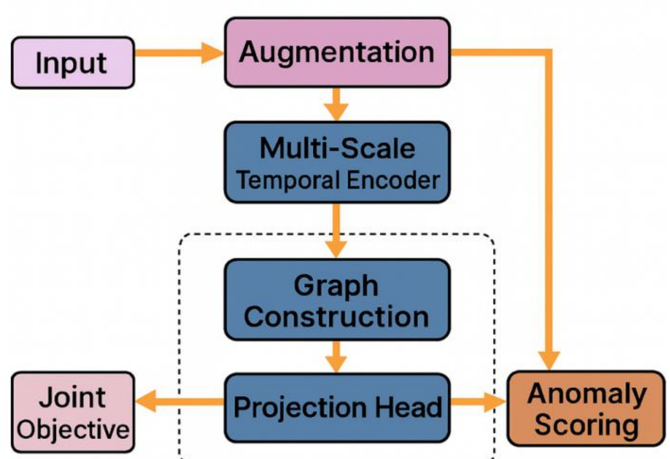

Figure 1. Overall model architecture

First, given a multidimensional indicator sequence $X = \{x_1, x_2, ..., x_T\}$ of the backend system, the model maps the original input to the implicit space representation h_t through the encoder f_θ. Its basic transformation is defined as:

$$h_t = f_\theta(x_t) = \mathrm{Re}LU(W_e x_t + b_e) \tag{1}$$

W_e and b_e are learnable parameters, and ReLU activation is used to preserve nonlinear features and suppress noisy responses. This process achieves a unified low-dimensional embedding of multidimensional time series signals, providing a feature foundation for subsequent dependency modeling and comparative optimization.

In the process of capturing dynamic dependencies between services, the model introduces an adaptive adjacency matrix A to characterize the structural relationship between indicators. The dynamic update rule based on feature similarity is defined as:

$$A_{ij} = \frac{\exp(sim(h_i, h_j)/\tau)}{\sum_{k=1}^{N} \exp(sim(h_i, h_k)/\tau)} \tag{2}$$

Where $sim(h_i, h_j)$ represents the cosine similarity and τ is the temperature coefficient. This mechanism can adaptively reflect the semantic correlation between various indicators under the system operation state, thereby dynamically capturing cross-service dependency changes.

In the feature interaction stage, in order to fuse the multi-layer information of the time dimension and the structure dimension, the spatiotemporal aggregation operation is defined as follows:

$$z_i = \sigma\left(\sum_{j=1}^{N} A_{ij} W_g h_j + W_t + h_i\right) \tag{3}$$

Where W_g and W_t are the weight matrices of the structural and temporal channels, respectively, and $\sigma(\cdot)$ represents the nonlinear activation function. This formula realizes the joint update of local neighborhood features and their own temporal state, enhancing the model's expressive power under dynamic dependency conditions.

In the contrastive learning stage, in order to ensure that "semantically similar" samples are aggregated and "semantically different" samples are separated in the representation space, a standardized contrastive loss function is introduced:

$$L_{con} = -\log \frac{\exp(sim(z_i, z_i^+)/\tau)}{\exp(sim(z_i, z_i^+)/\tau) + \sum_{k=1}^{K} (sim(z_i, z_i^-)/\tau)} \tag{4}$$

Among them, z_i^+ is a positive sample belonging to the same category as z_i, and z_k^- is a negative sample from a different context. By maximizing the inter-class difference and minimizing the intra-class distance, the model achieves semantically consistent clustering and enhanced anomaly separability under unlabeled conditions.

Finally, to further constrain temporal consistency and suppress feature drift, the model designs a dynamic smoothing regularization term:

$$L_{smooth} = \frac{1}{T-1}\sum_{t=1}^{T-1}\left\|z_{t+1} - z_t\right\|_2^2 \tag{5}$$

This constraint term is designed to maintain the stability of implicit representations across consecutive time steps, preventing false detections caused by transient fluctuations or system noise. Based on the overall design, the model's optimization objective is formulated as a weighted summation of multiple joint constraints, enabling anomaly representation learning that balances global consistency and local sensitivity at the feature level. This provides a solid theoretical and structural foundation for achieving highly robust anomaly detection in backend systems.

3. Performance Evaluation

3.1 Dataset

This study employs the Cloud Workload Dataset for Scheduling Analysis as the data foundation for validating the proposed method. The dataset records resource usage and scheduling behaviors of multiple virtual machines and tasks within a cloud computing platform, including performance metrics such as CPU utilization, memory consumption, task start and end times, job queue length, and I/O activity. The data are stored in a time-series format with multidimensional indicators, capturing the dynamic characteristics and sudden fluctuations of resource usage under varying load conditions in cloud environments.

Each sample in the dataset consists of multiple metrics within a given time window, allowing the construction of continuous sliding-window samples for sequence model training. This setup effectively simulates abnormal states that occur when backend services experience load surges or resource bottlenecks, aligning well with the research objectives of backend service anomaly detection. The dataset's multidimensional structure and dynamic variability enable anomaly detection models to capture semantic correlations among metrics and adapt to interference from diverse service dependency scenarios.

Moreover, the Cloud Workload Dataset for Scheduling Analysis demonstrates strong applicability in terms of sample scale and dimensional richness. It contains both stable daily load cycles and high-load competition scenarios, allowing models to learn the intrinsic distribution of normal behaviors and the deviation patterns associated with anomalies. Using this dataset for contrastive representation learning-driven evaluation enables a comprehensive assessment of the model's representational capability and anomaly identification performance across multidimensional cloud workload time series.

3.2 Experimental Results

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

Table1. Comparative experimental results

Method	Precision	Recall	F1-score	AUC	ADR
DCdetector [8]	0.842	0.790	0.815	0.923	0.88
MST-GAT[9]	0.816	0.758	0.785	0.907	0.85

DDMT[10]	0.829	0.774	0.800	0.915	0.86
Ours	0.872	0.821	0.846	0.941	0.91

Overall, the proposed contrastive representation learning-based anomaly detection method for backend services outperforms mainstream baseline models across multiple metrics. Compared with other approaches, Ours achieves a Precision of 0.872, significantly higher than DCdetector (0.842) and MST-GAT (0.816), indicating that the proposed model can more accurately distinguish between normal and abnormal states while reducing false alarms. This improvement primarily results from the model's ability to perform multi-scale semantic representation within the contrastive learning framework, which enhances the separability of inter-service dependencies in the embedding space and strengthens both detection accuracy and robustness.

In terms of Recall, Ours reaches 0.821, showing a clear advantage over other models in the 0.758-0.790 range. This suggests that the proposed method achieves broader anomaly coverage. The improvement arises from the integration of multi-scale feature encoding and dynamic adjacency modeling, allowing the model to perceive potential temporal and cross-service anomaly correlations. This enables the effective capture of propagative and collaborative anomalies within complex systems. Such capability is particularly important in backend service environments, where anomalies often spread in a cascading manner and single-point features fail to identify root causes effectively.

For the comprehensive metrics, F1-score and AUC, Ours achieves 0.846 and 0.941, respectively, demonstrating consistent and balanced detection performance. The improved F1-score indicates that the model maintains a favorable balance between precision and recall, while the high AUC value reflects stable discrimination across varying thresholds. These results validate the effectiveness of the semantic consistency constraint under the unsupervised contrastive learning framework, which reinforces the structural separability of the representation space. Consequently, the model remains both sensitive and generalizable to diverse anomaly patterns.

In terms of ADR (Anomaly Detection Rate), Ours achieves 0.91, the highest among all models, further confirming its capability to comprehensively capture abnormal distributions. Compared with reconstruction-based and graph-attention-based approaches, the proposed method jointly optimizes the contrastive loss and temporal consistency constraint, ensuring that implicit representations remain stable within dynamically evolving metric spaces. This enhances both detection continuity and global coherence. The results demonstrate that the proposed contrastive representation learning framework not only builds robust relationships among multidimensional features but also adaptively captures the complex evolution of backend system states, providing an effective solution for high-precision anomaly detection in intelligent operations.

This paper also evaluates the data sensitivity of pseudo-label noise and weak supervision ratio changes to contrast boundary consistency, and the experimental results are shown in Figure 2.

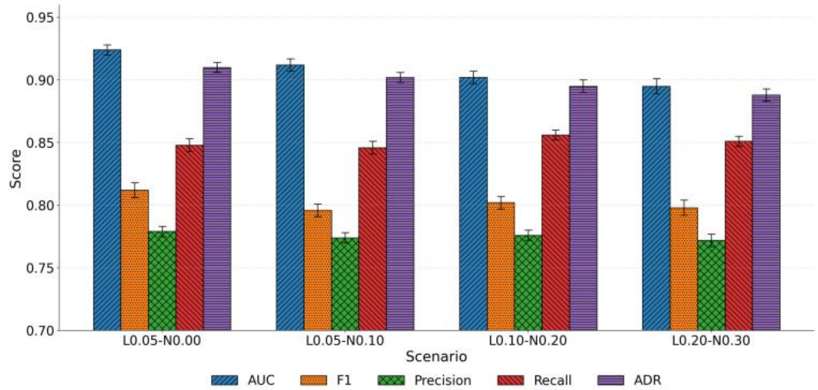

Figure 2. Experimental data sensitivity analysis of pseudo-label noise and weak supervision ratio changes on contrast boundary consistency

From the overall trend, as the ratio of pseudo-label noise increases and the weak supervision rate varies, the model's overall discriminative capability shows a gradual decline. The AUC decreases from 0.924 to 0.895, indicating that the boundary clarity of the contrastive representation space is affected when label quality deteriorates. Errors in pseudo-labels disrupt the semantic consistency between positive and negative sample pairs, preventing the model from maintaining clear separability between normal and abnormal samples at the feature distribution level. Consequently, the separability of potential anomalies in the projection space weakens. This trend reflects the sensitivity of contrastive learning to data consistency under weak supervision and highlights the importance of maintaining label confidence for effective anomaly identification in backend systems.

The variation in F1-score exhibits a "decline-slight recovery-decline" pattern (0.812 → 0.796 → 0.802 → 0.798). This suggests that moderate noise slightly perturbs the contrastive learning objective, while the diversity introduced by pseudo-labels provides a temporary data coverage effect, improving the model's generalization near boundary samples. This phenomenon implies that a moderate level of pseudo-label noise can help smooth the feature distribution, enhancing robustness near anomaly boundaries. However, as the noise ratio continues to rise, semantic discrimination between positive and negative samples becomes diluted, feature aggregation becomes blurred, and overall performance eventually declines.

The Precision metric remains relatively stable, fluctuating within a narrow range of 0.779-0.772, while Recall increases from 0.848 to 0.856 and slightly drops to 0.851, demonstrating consistent detection sensitivity. This trend indicates that the proposed multi-scale semantic encoding and dynamic adjacency modeling mechanisms can preserve anomaly detection capability under noisy conditions, avoiding excessive reliance on local label information. Even when pseudo-label errors occur, the model maintains continuous awareness of anomaly propagation patterns through temporal consistency constraints and structural self-calibration. This property validates the robustness of the proposed model in complex dynamic systems, where it sustains high recall and low volatility even under non-stationary distributions and metric drift.

The ADR (Anomaly Detection Rate) exhibits a gradual decline (0.910 to 0.888) with increasing noise and weak supervision ratios, indicating a moderate reduction in anomaly coverage. This occurs because pseudo-label uncertainty causes the model to adopt a more conservative confidence estimation for boundary samples, reducing the detection of marginal anomalies. Although the overall detection rate decreases, the reduction remains limited, demonstrating that the proposed method incorporates a strong self-repair mechanism through feature aggregation and semantic contrast, which offsets part of the disturbance introduced by pseudo-label noise. This result confirms the adaptability and fault tolerance of the contrastive learning framework in backend service anomaly detection, laying a solid foundation for future applications in more complex and low-supervision scenarios.

4. Conclusion

This study addresses the key challenges of anomaly detection in backend service environments, including complex dependencies, dynamic evolution, and label scarcity. It proposes an unsupervised detection method based on contrastive representation learning to achieve semantic-consistent modeling and enhanced anomaly separability in multidimensional time-series data. Through the joint design of multi-scale feature encoding, dynamic graph dependency modeling, and temporal consistency constraints, the model effectively captures latent inter-service relationships under unlabeled or weakly supervised conditions. This enables the formation of stable semantic aggregation and anomaly separation mechanisms in the feature space. Experimental results demonstrate that the proposed method exhibits strong generalization and robustness across multi-metric systems, maintaining high-precision anomaly recognition in complex, non-stationary environments and providing a new theoretical and algorithmic foundation for intelligent backend operations.

From a methodological perspective, the contrastive representation learning paradigm explored in this study offers a new viewpoint for traditional anomaly detection tasks. Unlike methods relying on explicit labels or reconstruction errors, the proposed approach extracts semantic relationships among features through self-supervised contrastive objectives, allowing the model to learn transferable anomaly discrimination boundaries in high-dimensional spaces. This mechanism not only reduces dependence on large-scale labeled data but also enhances model stability under non-stationary distributions, heterogeneous metrics, and noisy disturbances. Its applicability to cloud-native systems, microservice architectures, and intelligent operation scenarios suggests that contrastive learning can be effectively extended to broader domains such as complex system analysis, temporal diagnostics, and cross-domain anomaly identification.

At the application level, the significance of this research lies not only in performance improvement but also in advancing automation and interpretability in intelligent backend management. By incorporating structure-aware and semantic-consistency modeling, the system can adaptively recognize anomalies and reconstruct feature spaces during runtime, enabling early warnings when performance degradation, resource imbalance, or dependency anomalies occur. This mechanism provides a reliable foundation for key applications such as AIOps, automated resource scheduling, and cloud service quality assurance, while also guiding the development of self-healing cloud-native systems. As system scale continues to grow and operational complexity increases, such self-supervised detection methods will become a core technological support for maintaining the stability of cloud computing ecosystems.

Future research can be extended in three directions. First, introducing cross-domain contrastive mechanisms and hierarchical dependency modeling can improve adaptability to multi-tenant systems and heterogeneous metrics. Second, integrating generative modeling with causal inference frameworks can enable interpretable identification of anomaly causal chains. Third, exploring lightweight and distributed deployment strategies will allow real-time execution on edge nodes and hybrid cloud environments. Through deep integration with observability platforms, federated learning frameworks, and large-model inference systems, future contrastive representation learning-based anomaly detection methods are expected to play a greater role in intelligent monitoring, cloud security analysis, and service elasticity management, driving backend intelligent systems toward higher levels of autonomous decision-making and self-optimization.

References

[1] Y. Hundman, V. Constantinou, C. Laporte, I. Colwell and T. Soderstrom, "Detecting spacecraft anomalies using LSTMs and nonparametric dynamic thresholding," Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 387-395, 2018.

[2] H. Xu, W. Chen, N. Zhao, Z. Li, J. Bu, Z. Li, Y. Zheng and Z. Wu, "Unsupervised anomaly detection via variational auto-encoder for seasonal time series," Proceedings of the 35th International Conference on Machine Learning, pp. 4121-4130, 2018.

[3] Y. Su, Y. Zhao, C. Niu, R. Liu, W. Sun and D. Pei, "Robust anomaly detection for multivariate time series through stochastic recurrent neural network," Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 2828-2837, 2019.

[4] P. Malhotra, A. Ramakrishnan, G. Anand, L. Vig, P. Agarwal and G. Shroff, "LSTM-based encoder-decoder for multi-sensor anomaly detection," Proceedings of the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, 2016.

[5] H. Ren, B. Xu, Y. Wang, C. Yi, C. Huang, X. Kou, T. Xing, M. Yang, J. Tong and Q. Zhang, "Time-series anomaly detection service at Microsoft," Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 3009-3017, 2019.

[6] A. Deng and B. Hooi, "Graph neural network-based anomaly detection in multivariate time series," Proceedings of the AAAI Conference on Artificial Intelligence, pp. 4027-4035, 2021.

[7] J. Audibert, P. Michiardi, F. Guyard, S. Marti and M. Zuluaga, "USAD: Unsupervised anomaly detection on multivariate time series," Proceedings of the ACM SIGKDD Workshop on Outlier Detection and Description, 2020.

[8] C. Zhou and R. C. Paffenroth, "Anomaly detection with robust deep autoencoders," Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 665-674, 2017.

[9] G. Lai, W. C. Chang, Y. Yang and H. Liu, "Modeling long- and short-term temporal patterns with deep neural networks," Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 95-104, 2018.

[10] B. Zong, Q. Song, M. R. Min, W. Cheng, C. Lumezanu, D. Cho and H. Chen, "Deep autoencoding Gaussian mixture model for unsupervised anomaly detection," Proceedings of the International Conference on Learning Representations, 2018.