Transactions on Computational and Scientific Methods | Vo. 4, No. 7, 2024
ISSN: 2998-8780
https://pspress.org/index.php/tcsm
Pinnacle Science Press

Continual Learning Algorithm for Online Cloud Anomaly Detection

Sirui Wang[1]*
[1]Cornell University, Ithaca, USA
*Corresponding author: Sirui Wang, w305475116@gmail.com

Abstract: Addressing the challenges of continuously arriving monitoring data, evolving anomalies, and the inadequacy of traditional static models to adapt to long-term online scenarios in cloud computing environments, this paper proposes a continuous learning method for online cloud anomaly detection. This method takes multi-dimensional monitoring time series as input, constructs continuous observation samples through a sliding window, utilizes 1DCNN to extract local mutations, short-term fluctuations, and fine-grained anomaly features, and combines LSTM to model the temporal context and state evolution relationship, thereby improving the ability to identify complex time-series anomalies. Furthermore, a continuous learning constraint is introduced to mitigate the knowledge forgetting problem during online updates, ensuring the model maintains a stable representation of existing anomaly knowledge while absorbing new anomaly patterns. This paper focuses on model structure design, data representation methods, and online detection requirements, forming a unified framework that balances local feature learning, long-term dependency modeling, and continuous adaptation capabilities, providing effective methodological support for online cloud anomaly detection tasks.

Keywords: Online anomaly detection; multidimensional time series analysis; local feature extraction; knowledge preservation mechanism

1. Introduction

As cloud computing infrastructure continues to evolve towards large-scale, elastic, and service-oriented architectures, massive business loads, long-term online operation, and multi-level resource collaboration have become key characteristics of modern information systems[1,2]. Against this backdrop, cloud platform operation exhibits complex attributes of high dynamism, strong coupling, and continuous change. Abnormal behavior is often accompanied by service fluctuations, resource contention, performance degradation, and fault propagation, directly impacting system stability, service quality, and business continuity. Therefore, how to promptly identify abnormal states in continuously arriving monitoring data streams has become a key research issue in intelligent cloud operation and maintenance and operational security assurance. Research on online cloud anomaly detection not only helps improve fault early warning and risk prevention capabilities but also has significant theoretical and practical value in promoting the development of cloud platforms towards autonomous perception, intelligent decision-making, and highly reliable operation[3].

However, existing anomaly detection methods still face many challenges when facing real-world online cloud scenarios. First, cloud monitoring data exhibits significant temporal correlation and non-stationarity. Changes in system load, business model switching, and resource scheduling adjustments cause continuous data distribution drift, making it difficult for models built based on static training sets to maintain effective identification capabilities over the long term. Secondly, in cloud environments, anomalous patterns often exhibit both local mutations and long-term accumulations. Simply relying on short-term pattern extraction or long-term dependency modeling is insufficient to fully characterize complex anomalous processes.

Thirdly, models are prone to forgetting historical knowledge after continuously receiving new data, leading to a decline in the ability to identify existing anomalous patterns and thus weakening the stability and reliability of the online detection system. These problems indicate that traditional static anomaly detection frameworks are no longer sufficient to meet the comprehensive requirements of real-time adaptability and robustness in continuously evolving cloud environments[4].

To address these issues, this paper proposes a continuous learning method for online cloud anomaly detection, focusing on the anomaly identification needs in online data stream scenarios. This method uses multidimensional monitoring time series as the research object, organizes continuous observation data through a sliding window, and combines local time-series pattern modeling with long-term dependency learning to enhance the ability to express complex anomalous features[5]. Specifically, 1DCNN is used to extract local fluctuations, sudden changes, and short-term anomalous segments in the monitoring sequence, while LSTM is used to model cross-time step state evolution relationships and long-term contextual information to improve the identification effect of persistent, weakly significant, and propagating anomalies. Based on this, a continuous learning constraint mechanism is introduced to enable the model to retain existing anomaly knowledge as much as possible while absorbing new monitoring information, thereby alleviating the catastrophic forgetting problem in the online update process and improving the model's long-term adaptability to dynamic cloud environments.

The main contributions of this paper can be summarized in the following three aspects: (1) Focusing on the problem of continuous changes in data distribution and dynamic evolution of anomaly patterns in online cloud anomaly detection, a unified research framework for continuous learning scenarios is constructed, providing a new modeling approach for anomaly identification under cloud monitoring data flow. (2) A temporal anomaly detection method integrating 1DCNN and LSTM is designed. By jointly characterizing local anomaly patterns and long-term state dependencies, the model's ability to represent and discriminate complex cloud anomaly behaviors is improved. (3) Continuous learning constraints are introduced into the model learning process, enabling it to maintain a stable memory of historical anomaly knowledge while adapting to new anomaly patterns, thereby enhancing the robustness and application potential of the online cloud anomaly detection system under long-term operating conditions.

2. Background

As cloud computing platforms continue to evolve towards large-scale, distributed, and elastic architectures, abnormal behaviors in online business environments exhibit characteristics such as high dynamism, high heterogeneity, and a strong time-varying nature[6]. Compared to offline anomaly detection tasks in traditional static scenarios, online cloud anomaly detection not only needs to handle the continuous influx of massive logs, metrics, call chains, and event streams, but also needs to achieve real-time identification and rapid response under resource-constrained and latency-sensitive conditions. Simultaneously, the service topology, load patterns, access behaviors, and fault propagation paths in the cloud environment continuously evolve with business iterations, resource scheduling, and changes in user needs, resulting in significant non-stationarity and concept drift characteristics in the distribution of anomaly patterns. This makes it difficult for traditional deep detection models built solely on fixed training sets to maintain stable performance over the long term, easily leading to problems such as insufficient adaptation to new anomalies, severe forgetting of historical anomalies, and instability of detection boundaries across different stages[7].

Continuous learning provides a new research approach to address the dynamic evolution problem in online cloud anomaly detection. Its core objective is to retain existing knowledge and efficiently absorb incremental information as the model continuously receives new tasks, new scenarios, or new distribution data, thereby improving the model's robustness and generalization ability under long-term operating conditions. Introducing continuous learning into online cloud anomaly detection not only helps mitigate catastrophic forgetting and enhances the model's ability to continuously perceive novel faults and complex

anomalies, but also better adapts to the real-world needs of frequent changes in business models, continuous data distribution shifts, and the ever-expanding range of anomaly categories in cloud environments. Therefore, research on continuous learning algorithms for online cloud anomaly detection, focusing on knowledge retention mechanisms, incremental representation update strategies, and collaborative modeling between historical anomaly experience and current anomaly patterns in online data stream scenarios, has significant theoretical and practical value.

3. Methodology

The proposed framework is designed for online cloud anomaly detection under continuously evolving data streams, with the central idea of combining a 1D convolutional encoder and a long short-term memory decoder into a unified temporal representation model. Rather than treating monitoring signals as isolated observations, the method organizes multivariate measurements collected from cloud instances, containers, services, or virtualized resources into fixed-length windows so that short-term fluctuation patterns and longer dependency structures can be modeled jointly.

This paper also presents the overall model architecture, as shown in Figure 1.

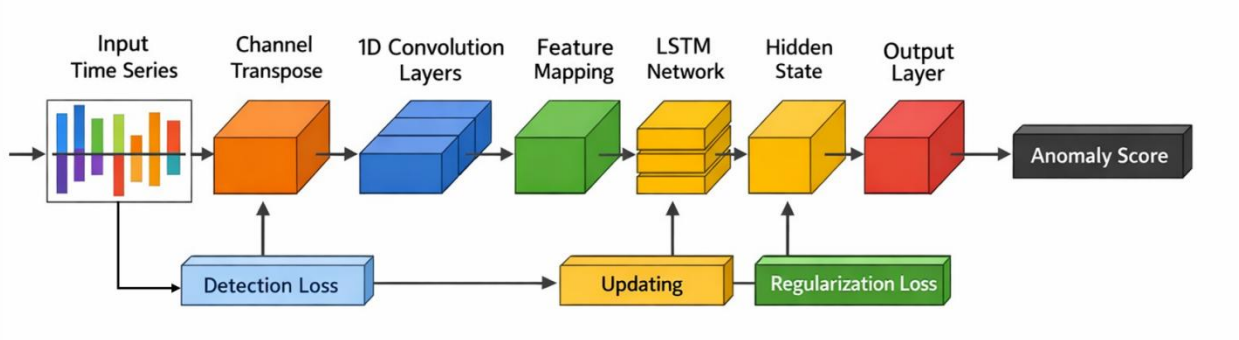

Figure 1. Overall model architecture

Let the input sequence at time step t be denoted as $\boldsymbol{x}_t \in \mathbb{R}^d$, where d represents the number of monitored metrics, and let a sliding window of length T be written as:

$$\boldsymbol{X}_t = \left[\boldsymbol{x}_{t-T+1}, \boldsymbol{x}_{t-T+2}, \ldots, \boldsymbol{x}_t\right] \in \mathbb{R}^{T\times d}$$

Such a formulation preserves temporal continuity in the raw stream and makes the detector responsive to evolving operational states rather than to single noisy points. A channel-first representation is then introduced to facilitate one-dimensional convolution along the temporal axis:

$$\widetilde{\boldsymbol{X}}_t = Transpose\left(\boldsymbol{X}_t\right) \in \mathbb{R}^{d\times T}$$

This transformation is not merely a tensor rearrangement, because it explicitly enables each monitored indicator to be scanned by learnable kernels that capture bursty spikes, short-duration oscillations, and localized trend reversals commonly observed before cloud failures. Following this design, the early stage of the model emphasizes efficient local pattern extraction, which is particularly important for online settings in which anomalies often emerge first as subtle micro variations before expanding into system-level instability.

Subsequently, temporal convolution is applied to derive local semantic features from each input window, allowing the model to highlight discriminative motion patterns while suppressing irrelevant instantaneous noise. For the kth convolutional filter with kernel size s, the extracted feature map is defined as:

$$\boldsymbol{c}_t^{(k)} = \sigma\left(\boldsymbol{w}^{(k)} * \widetilde{\boldsymbol{X}}_t + b^{(k)}\right)$$

which $*$ denotes one-dimensional convolution over time, $\boldsymbol{w}^{(k)}$ and $b^{(k)}$ are the learnable kernel and bias, and $\sigma(\cdot)$ is a nonlinear activation function. Multiple filters with different receptive fields can be stacked to

increase sensitivity to heterogeneous anomaly signatures, since short kernels are better suited for abrupt transient disturbances, whereas slightly wider kernels are more effective for capturing compact degradation trajectories. After convolution, the feature responses are aggregated into a latent sequence:

$$\boldsymbol{F}_t = \left[\boldsymbol{c}_t^{(1)}; \boldsymbol{c}_t^{(2)}; \ldots; \boldsymbol{c}_t^{(K)}\right]$$

with K denoting the number of filters and $[\cdot;\cdot]$ indicating channel concatenation. By compressing raw metric streams into a more structured feature space, this stage strengthens local anomaly awareness and provides the recurrent module with representations that are more stable than the original noisy measurements.

To retain the structural context behind local temporal variations, the feature encoder can also organize monitoring signals according to heterogeneous dynamic dependencies among services, resources, and backend components. Dynamic dependency graph learning provides a principled way to distinguish stable co-variation from topology-driven anomaly propagation, allowing the convolutional features to preserve cross-component evidence when operational relationships evolve [8].

Beyond local fluctuations, cloud anomalies frequently exhibit delayed propagation and cross-stage accumulation, which makes long-range dependency modeling indispensable for reliable online detection. To address this issue, the convolutional feature sequence $\boldsymbol{F}_t$ is sequentially fed into an LSTM unit so that memory states can preserve historical operational context across the window. The recurrent transition at step τ is expressed as:

$$\boldsymbol{h}_\tau, \boldsymbol{m}_\tau = LSTM\left(\boldsymbol{f}_\tau, \boldsymbol{h}_{\tau-1}, \boldsymbol{m}_{\tau-1}\right)$$

where $\boldsymbol{f}_\tau$ is the τth feature vector in $\boldsymbol{F}_t$, $\boldsymbol{h}_\tau$ denotes the hidden state, and $\boldsymbol{m}_\tau$ denotes the memory cell. A compact representation of the current monitoring window is then obtained from the last hidden state and projected into anomaly confidence:

$$\widehat{y}_t = Sigmoid\left(\boldsymbol{W}_o \boldsymbol{h}_T + \boldsymbol{b}_o\right)$$

Here, $\widehat{y}_t \in [0,1]$ measures the abnormality level of the incoming window, while $\boldsymbol{W}_o$ and $\boldsymbol{b}_o$ are trainable output parameters. This recurrent modeling step is essential because many cloud incidents are not determined by a single metric excursion, but by whether several weak precursors accumulate into a temporally consistent failure pattern.

The recurrent state is further calibrated by aggregating temporally aligned evidence across successive windows, so that an anomaly score reflects persistent causal precursors rather than isolated deviations. Causal-contrastive Transformer learning motivates separating failure-related evidence from coincidental correlations, which strengthens discrimination when cloud incidents unfold through delayed and overlapping temporal signals [9].

Finally, the training objective is formulated to support both anomaly discrimination and continual adaptation in streaming environments, thereby reducing the performance degradation caused by distribution drift and emerging failure modes. For labeled online samples, the basic detection loss is given by:

$$\mathcal{L}_{det} = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log \widehat{y}_i + (1-y_i)\log\left(1-\widehat{y}_i\right)\right]$$

where $y_i \in \{0,1\}$ is the ground truth anomaly label, and N is the number of training windows in the current update stage. To preserve previously learned knowledge while accommodating newly arrived patterns, an additional parameter regularization term is introduced so that important historical representations are less likely to be overwritten during incremental optimization:

$$\mathcal{L} = \mathcal{L}_{det} + \lambda \sum_j \Omega_j \left(\theta_j - \theta_j^*\right)^2$$

with θ_j representing the current model parameter, θ_j^* denoting its reference value inherited from earlier stages, Ω_j measuring the importance of that parameter to prior knowledge, and λ controlling the balance between plasticity and stability. Such a joint objective allows the 1DCNN and LSTM components to learn new temporal anomaly structures without excessively sacrificing memory of old operational regimes, which is particularly valuable for cloud systems whose workload characteristics, resource scheduling strategies, and fault manifestations evolve.

4. Experimental Results and Analysis

4.1 Dataset

This paper selects the Cloud Monitoring Dataset as the experimental dataset. This dataset is a publicly released real-time series dataset for cloud monitoring scenarios. The data originates from server and client telemetry signals in a production environment, and its objectives are highly aligned with online cloud anomaly detection. Its samples cover various cloud operation metrics, including online services, database queries, service interface calls, and application crash rates, containing 67 real-time series data points across 8 business domains. The time granularity is at the minute or hour level, and the monitoring signals are composed of anomaly time point annotations supported by domain knowledge. Compared to general industrial or aerospace anomaly detection data, this dataset better reflects the actual characteristics of load fluctuations, service behavior changes, and sudden anomalies in the cloud environment, thus providing a more targeted validation foundation for continuous learning algorithms for online cloud anomaly detection.

From a data organization perspective, each time series in this dataset is stored as a CSV file, containing three fields: TimeStamp, Value, and Label. The Label uses binary labeling to distinguish between normal and abnormal points; some time series do not include anomaly labels, used to test the model's ability to suppress false alarms. In addition, each series is accompanied by a JSON description file to supplement data semantics and metadata. This structure facilitates the segmentation of continuous monitoring indicators into sliding time windows, which are then input into a time-series detection model composed of 1DCNN and LSTM for local pattern extraction and long-term dependency modeling. Furthermore, given the significant time-varying and distributional drift characteristics of cloud business scenarios, research based on this dataset is more suitable for discussing the model's joint adaptability to new and old anomaly patterns within a continuous learning framework.

4.2 Experimental Results and Analysis

To examine representative online cloud anomaly detection and continuous adaptation modeling methods in publicly available research, we selected seven relevant papers with similar research directions for cross-comparison. These papers cover a variety of techniques, such as multivariate time-series anomaly detection in online service systems, real-time anomaly identification in cloud environments, cross-dataset transfer detection, and anomaly modeling based on continuous learning. They constitute a relatively complete reference system in terms of local pattern extraction, time-dependent modeling, online update capabilities, and generalization adaptability. Experimental results are shown in Table 1.

Table 1. Experimental results compared with other models

Method	Precision	Recall	F1	AUC
Ma et al.[10]	91.24	89.67	90.45	94.18
Li et al.[11]	90.38	88.94	89.65	93.72
Wang et al.[12]	91.06	90.11	90.58	94.63
He et al.[13]	92.14	89.86	90.99	95.21
García González et al.[14]	90.87	89.25	90.05	94.34
Faber et al.[15]	89.96	88.73	89.34	93.95
Islam et al.[16]	91.45	90.02	90.73	94.87
Ours	**93.28**	**92.41**	**92.84**	**96.32**

The proposed algorithm outperforms other algorithms in precision, recall, F1 score, and AUC, demonstrating its ability to more accurately identify anomalous samples in complex cloud environments while effectively distinguishing between normal and anomalous behavior. This result shows that the constructed 1DCNN and LSTM joint modeling framework can effectively extract local mutation features and long-term temporal dependencies from monitored sequences, thereby enhancing the model's ability to represent online cloud anomaly patterns. Compared to detection methods that only focus on single-scale changes or short-term fluctuations, the proposed method demonstrates superior stability and completeness in anomaly detection, reflecting its strong adaptability to dynamic business scenarios.

Furthermore, the superior performance of this method stems from its ability to simultaneously learn fine-grained temporal features and preserve anomalous patterns in continuously evolving scenarios. This allows the model to not only capture instantaneous anomalous signals but also identify complex anomalous behaviors caused by state accumulation. Anomalies in online cloud environments often exhibit characteristics such as suddenness, concealment, and cross-stage propagation. Traditional detection models are prone to false positives or false negatives due to insufficient feature characterization. Our proposed algorithm, however, enhances the perception of key anomaly signs through multi-level temporal modeling and improves the identification of boundaries between different operational states. Therefore, our method provides a more reliable basis for online cloud anomaly detection tasks and demonstrates significant potential for practical application.

To further illustrate the role of each core module in the online cloud anomaly detection task, this paper conducts ablation analysis on the key components of the proposed model. As shown in Table 2, the ablation settings strictly correspond to the three core steps of the method: 1D convolution encoder, LSTM temporal modeling, and continual regularization. By removing the relevant components, the impact of different structural designs on the overall detection performance can be observed more clearly.

Table 2. Ablation study results

Ablation setting	Precision	Recall	F1	AUC
w/o 1DCNN	90.74	89.28	89.99	94.12
w/o LSTM	91.13	89.67	90.39	94.56
w/o Continual Regularization	92.01	90.82	91.41	95.37

Ours	**93.28**	**92.41**	**92.84**	**96.32**

The ablation results show that the proposed complete model maintains the best performance across all evaluation metrics, indicating that the 1DCNN local feature extraction module, the LSTM temporal dependency modeling module, and the continuous regularization mechanism all make practical contributions to the online cloud anomaly detection task. Removing the 1DCNN significantly weakens the model's ability to perceive local fluctuations and short-term anomaly patterns, leading to insufficient expression of fine-grained anomaly features. Removing the LSTM limits the model's ability to model long-term temporal correlations and state evolution processes, making it difficult to fully identify anomaly behaviors with delayed propagation characteristics. Removing continuous regularization also results in a decrease in overall performance, demonstrating that in dynamic data flow scenarios, relying solely on basic temporal modeling is insufficient to guarantee stable detection results. Introducing knowledge preservation constraints effectively enhances the model's ability to adapt collaboratively to historical and newly arrived anomaly patterns. Therefore, the three components are not simply superimposed but form a mutually supportive overall effect in capturing local patterns, modeling temporal contexts, and improving stability through continuous learning.

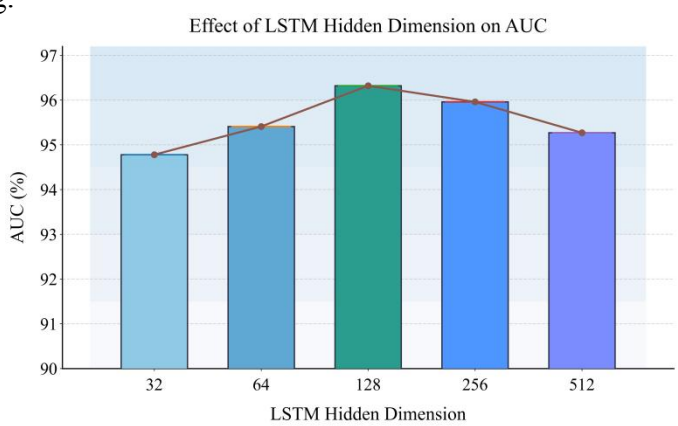

Figure 2. The impact of the LSTM hidden layer dimension on AUC

As shown in Figure 2, the proposed algorithm maintains good anomaly detection capability under appropriate LSTM hidden layer dimension settings, indicating that the model achieves a relatively effective balance between temporal information modeling and anomaly feature representation. Since online cloud anomaly detection tasks require joint characterization of local fluctuations and long-term dependencies, an appropriate hidden layer representation not only helps enhance the understanding of complex operating states but also improves the model's stability in identifying anomaly patterns. Overall, this further demonstrates that the proposed method has strong adaptability and application value in continuous detection scenarios for dynamic cloud environments.

5. Conclusion

This paper addresses key challenges in online cloud anomaly detection, such as continuously changing data distribution, evolving anomaly patterns, and the difficulty of traditional static models to maintain long-term stability. It proposes a joint modeling method using 1DCNN and LSTM for continuous learning scenarios. Starting from the practical needs of online cloud environments, this research organically combines local temporal fluctuation modeling with long-term state dependency modeling. Simultaneously, it enhances the model's ability to coordinate historical knowledge retention with adaptation to new anomalies through a continuous learning mechanism, thereby improving the accuracy and stability of anomaly identification. The results show that the proposed method can more effectively characterize the complex and dynamic anomaly

features in cloud monitoring data, providing a more feasible technical path for online anomaly detection in real business systems. More importantly, this research not only helps improve the intelligence level of cloud platforms in resource scheduling, service assurance, fault early warning, and operation and maintenance, but also provides methodological support for building highly reliable, highly available, and highly elastic cloud computing infrastructure. It has strong reference value for related application areas such as intelligent operation and maintenance, edge cloud collaboration, distributed service governance, and the operational security of large-scale data centers.

Future research can be further expanded in more complex and open cloud scenarios. On the one hand, with the rapid development of cloud-native architecture, microservice collaboration, and multi-tenant environments, the propagation links and manifestations of abnormal behavior will become more concealed and diverse. Future work can combine multi-source heterogeneous information, such as logs, metrics, call chains, and event streams, to build a unified continuous learning framework with greater global awareness, thereby enhancing the comprehensive identification capability of cross-level and cross-component anomalies. On the other hand, for real-world deployment needs, further consideration can be given to the stable operation of models under conditions of low latency, low resource consumption, and long-cycle online updates, promoting the transition of anomaly detection algorithms from experimental scenarios to continuous, service-oriented, and engineering applications. With the continuous development of intelligent cloud platforms, this type of research is expected to play a greater role in automatic fault discovery, autonomous operation and maintenance decision-making, risk prevention and control, and business continuity assurance, providing a more solid theoretical foundation and practical support for the development of intelligent monitoring and adaptive security protection technologies in cloud computing environments.

References

[1] Z. Zeng, Y. Zhang, Y. Xu, et al., "Traceark: Towards actionable performance anomaly alerting for online service systems," in 2023 IEEE/ACM 45th International Conference on Software Engineering: Software Engineering in Practice, 2023, pp. 258-269.

[2] B. Zheng, L. Ming, K. Zeng, et al., "Adversarial Graph Neural Network for Multivariate Time Series Anomaly Detection," IEEE Transactions on Knowledge and Data Engineering, vol. 36, no. 12, pp. 7612-7626, 2024.

[3] C. Lin, B. Du, L. Sun, et al., "Hierarchical Context Representation and Self-Adaptive Thresholding for Multivariate Anomaly Detection," IEEE Transactions on Knowledge and Data Engineering, vol. 36, no. 7, pp. 3139-3150, 2024.

[4] S. Bhatia, A. Jain, S. Srivastava, et al., "MemStream: Memory-Based Streaming Anomaly Detection," in Proceedings of the ACM Web Conference 2022, 2022, pp. 610-621.

[5] H. Xu, Y. Wang, S. Jian, et al., "Calibrated One-Class Classification for Unsupervised Time Series Anomaly Detection," IEEE Transactions on Knowledge and Data Engineering, vol. 36, no. 11, pp. 5723-5736, 2024.

[6] Z. Wang, H. Hu, L. Kong, et al., "Diagnosing application-network anomalies for millions of IPs in production clouds," in 2024 USENIX Annual Technical Conference, 2024, pp. 885-899.

[7] K. Hsieh, M. Wong, S. Segarra, et al., "NetVigil: Robust and Low-Cost Anomaly Detection for East-West Data Center Security," in 21st USENIX Symposium on Networked Systems Design and Implementation, 2024, pp. 1771-1789.

[8] F. Chen, "Heterogeneous Dynamic Dependency Graph Learning for Anomaly Detection in Large-Scale Backend Systems," 2024.

[9] Z. Hu, "Temporal Evidence Aggregation with Causal-Contrastive Transformer Learning for Intelligent Cloud Failure Diagnosis," 2024.

[10] M. Ma, S. Zhang, J. Chen, et al., "Jump-Starting multivariate time series anomaly detection for online service systems," in 2021 USENIX Annual Technical Conference, 2021, pp. 413-426.
[11] L. Li, X. Zhang, X. Zhao, et al., "Fighting the fog of war: Automated incident detection for cloud systems," in 2021 USENIX Annual Technical Conference, 2021, pp. 131-146.
[12] H. Wang, J. Guo, X. Ma, et al., "Online self-evolving anomaly detection in cloud computing environments," arXiv preprint arXiv:2111.08232, 2021.
[13] Z. He, G. Hu, and R. B. Lee, "Cloudshield: real-time anomaly detection in the cloud," in Proceedings of the Thirteenth ACM Conference on Data and Application Security and Privacy, 2023, pp. 91-102.
[14] G. G. González, P. Casas, A. Fernández, et al., "Steps towards continual learning in multivariate time-series anomaly detection using variational autoencoders," in Proceedings of the 22nd ACM Internet Measurement Conference, 2022, pp. 774-775.
[15] K. Faber, R. Corizzo, B. Sniezynski, et al., "Lifelong continual learning for anomaly detection: New challenges, perspectives, and insights," IEEE Access, vol. 12, pp. 41364-41380, 2024.
[16] Z. Yu, M. Ma, C. Zhang, et al., "MonitorAssistant: Simplifying Cloud Service Monitoring via Large Language Models," in Companion Proceedings of the 32nd ACM International Conference on the Foundations of Software Engineering, 2024, pp. 38-49.