Transactions on Computational and Scientific Methods | Vo. 4, No. 12, 2024
ISSN: 2998-8780
https://pspress.org/index.php/tcsm
Pinnacle Science Press

Federated Multi-Scale Representation Learning for Privacy-Aware Log Anomaly Detection in Distributed Cloud Environments

Zhijun Wang
Rice University, Houston, USA
g674901087@gmail.com

Abstract: This work addresses scenarios where multiple organizations operate cloud native systems and face sensitive log data, restricted sharing, and cross-domain anomaly propagation that is difficult to characterize. It proposes a federated representation-based framework for log anomaly identification. Each participant structures and semantically encodes its local log sequence to produce low-dimensional latent features, and dynamic dependencies of system behavior are captured through multi-scale temporal aggregation with sliding windows. In cross-domain collaboration, federated updates and parameter aggregation enable knowledge sharing without exposing raw logs, while a representation-level consistency constraint mitigates distribution differences and semantic drift across clients, improving stability and transferability of the shared subspace. For anomaly modeling, the method constructs a distance-driven local risk scoring strategy using the normal semantic center as reference, providing interpretable abnormal discrimination even under weak or missing labels. Heterogeneous multi-client settings are built on a unified public log benchmark, and systematic sensitivity evaluation on optimizer choice, representation dimension, anomaly proportion, and template noise demonstrates stable performance and strong cross-domain robustness under diverse perturbations.

Keywords: Semantic consistency constraints, temporal multi-scale aggregation, distributed heterogeneous robustness, risk distance scoring

1. Introduction

Cloud native systems have increasingly complex business interactions, and the collaboration and sharing of data across organizations continue to expand. Logs serve as structured records of service operating states, link behaviors, and load evolution. They are an essential foundation for ensuring system stability and business security. However, logs usually contain highly sensitive content, such as access paths, request contexts, dependency features, and internal business processes. When logs are transmitted or aggregated across organizations for analysis, severe risks may arise, including privacy leakage, data misuse, and malicious correlation inference. These risks are more prominent in cross-organization settings because log sources are heterogeneous, formats vary, and sensitivity levels differ. Meanwhile, organizations have natural privacy boundaries. Information sharing is necessary, yet difficult to achieve. This situation highlights the importance of a mechanism that can protect privacy while enabling collaborative anomaly detection[1].

Traditional anomaly detection relies on centralized modeling. In cross-organization environments, this is often infeasible. First, transferring data to a central location creates severe privacy exposure. Second, differences in log structures make feature alignment difficult if logs are directly combined. Third,

anonymization such as desensitization, offline cleaning, or aggregation cannot fully prevent reverse inference. At the same time, distributed microservice architectures mean anomalies propagate along service chains and evolve across dimensions. Local logs alone cannot capture cross-system dynamic relations. If organizations cannot achieve joint modeling that is collaborative but privacy-preserving, anomaly recognition remains local and cannot reveal essential risks[2].

Unlike pure parameter aggregation, anomaly detection on logs depends more on semantic structures, invocation logic, and implicit behavioral distributions. Simple parameter averaging cannot support cross-domain collaborative analysis. A representation learning mechanism is needed. It should allow organizations to extract high-level semantic representations without exposing raw logs or sensitive features, and then align these representations collaboratively. This helps the model understand real behavioral patterns while avoiding data leakage. As anomalies are sparse and semantically ambiguous, rule-based or shallow statistical methods cannot describe them effectively. Deep representation learning becomes even more critical.

In this context, applying federated learning has strategic value. Its core advantage is that data remains local. Only model parameters or embeddings are shared. Centralized storage of sensitive logs is avoided. However, classical federated approaches face non-independent distributions, semantic fragmentation, and local minima. Cross-organization representations may be insufficient, and knowledge transfer may be unstable. A framework is needed that supports cross-domain alignment, semantic modeling for diverse logs, and collaborative understanding under privacy constraints. Organizations should maintain data sovereignty, compliance boundaries, and privacy completeness, while still forming common awareness of anomaly patterns. Through federated representation learning, an iterative flow of local learning, collaborative sharing, and representation aggregation can be achieved. This forms cross-organization risk cognition[3].

This direction has theoretical and practical value. Theoretically, it breaks the assumption that modeling requires centralized data. It explores how to achieve collaborative anomaly understanding when data distributions differ and privacy requirements are strict. It promotes the integration of federated learning, representation learning, and log anomaly detection in cross-organization environments. It builds a collaborative intelligence mechanism that is generalizable, portable, and scalable. It also provides new insights into training shared representations under distributed security constraints, preventing semantic leakage, and increasing anomaly separability at the representation level. Practically, it supports cross-institution supervision, cross-organization operation collaboration, link-level security monitoring, and distributed cloud risk control. It helps establish a more robust, agile, and transparent security ecosystem than any single organization can achieve alone.

In summary, cross-organization anomaly detection with privacy protection for cloud logs is driven by real needs and is a fundamental capability for collaborative governance in cloud ecosystems. Federated representation learning can enable shared anomaly representations without sacrificing privacy and compliance. It pushes cloud security from isolated defense to collaborative intelligence. It will have a lasting influence on next-generation cloud native security technologies.

2. Related work

Research on log-based anomaly detection has evolved significantly from early rule-driven and statistical approaches to modern deep learning-based paradigms. Initial studies focus on modeling sequential patterns in system logs using recurrent architectures and probabilistic assumptions. Representative works such as DeepLog and LogAnomaly demonstrate that sequential dependencies in logs can effectively characterize system behavior and detect deviations [4-5]. Subsequent efforts further address challenges such as log instability, noise, and parsing complexity, proposing robust detection mechanisms and preprocessing pipelines [6-8]. Meanwhile, log parsing techniques, including online parsing and automated template extraction, provide the structural foundation for downstream anomaly modeling [9-10]. In addition, time-series-based anomaly detection methods, particularly those leveraging LSTM architectures and dynamic

thresholding strategies, further enhance the ability to capture temporal irregularities in large-scale systems [11-12].

With the increasing complexity of distributed systems, anomaly detection has gradually shifted toward representation learning-based frameworks. Instead of relying on handcrafted features, these methods aim to learn compact latent representations that preserve semantic and structural information. Foundational work on representation learning establishes the theoretical basis for mapping high-dimensional observations into meaningful embedding spaces [13]. Building on this, one-class classification and deep anomaly detection methods enable effective identification of abnormal patterns under weak or even no supervision [14-15]. Recent surveys and empirical studies highlight that deep learning has become the dominant paradigm for log anomaly detection in both academia and industry, especially in large-scale distributed environments [16-18]. These methods significantly improve detection accuracy and generalization capability, but they typically assume centralized data availability, which limits their applicability in privacy-sensitive scenarios.

To overcome the limitations of centralized modeling, federated learning has emerged as a promising framework for collaborative anomaly detection. The seminal work on communication-efficient federated optimization provides the foundation for decentralized model training without sharing raw data [19]. Subsequent studies further explore theoretical challenges and system-level constraints in federated settings [20]. In the context of anomaly detection, federated approaches have been applied to IoT and distributed systems, demonstrating that collaborative learning can improve detection performance while preserving data locality [21-22]. More recent works incorporate graph structures and contrastive learning into federated frameworks, enabling more expressive modeling of system dependencies and cross-client relationships [23-24]. However, these approaches still face challenges related to data heterogeneity, representation inconsistency, and the difficulty of capturing shared semantic structures across organizations.

In parallel, advances in representation learning techniques, particularly contrastive learning, have further improved the robustness of anomaly detection models. Contrastive frameworks learn discriminative representations by maximizing agreement between similar samples while separating dissimilar ones, which has been shown effective in various domains [25-26]. Additionally, graph-based models, including graph convolutional networks and graph attention networks, provide powerful tools for modeling structural dependencies in complex systems [27-28]. These methods are especially relevant for microservice architectures and distributed environments, where anomalies often propagate through interconnected components rather than isolated events.

From a system perspective, large-scale cloud environments introduce additional challenges such as tail latency, resource contention, and service-level variability. Foundational studies on large-scale system behavior emphasize the importance of modeling rare but critical events that dominate system performance [29]. Recent work on AI-driven system optimization and microservice lifecycle management further highlights the need for intelligent anomaly detection mechanisms that can operate under dynamic and heterogeneous conditions [30]. In this context, anomaly detection is no longer an isolated task but an integral component of intelligent system management.

Recent studies also explore the integration of advanced AI techniques, including generative modeling, uncertainty estimation, and multi-scale temporal learning, into anomaly detection frameworks. These approaches aim to improve robustness and interpretability by modeling uncertainty and capturing hierarchical temporal patterns [31-34]. At the same time, combining time-series modeling with graph structures enables more comprehensive characterization of system dynamics, particularly in microservice environments where interactions are complex and evolving [35-36].

Beyond traditional anomaly detection, the rapid development of large language models and privacy-preserving learning techniques introduces new opportunities for intelligent system analysis. Methods for low-resource adaptation, structural regularization, and privacy-aware fine-tuning provide new tools for modeling complex data distributions without compromising sensitive information [37-40]. These techniques have been

applied to various domains, including financial risk analysis, distributed time-series modeling, and intelligent scheduling, demonstrating their potential for enhancing anomaly detection and decision-making systems [41-45].

In addition, a range of application-driven studies further extend anomaly detection to recommendation systems, user behavior modeling, and multi-task learning scenarios [46-49]. Emerging research on autonomous agents, logical reasoning, and multi-agent collaboration suggests that future anomaly detection systems may evolve toward more intelligent and adaptive frameworks capable of reasoning over complex system states [50-52].

3. Method

This method aims to achieve collaborative anomaly characterization modeling without cross-organizational data leaving its domain. First, each organization performs a structured representation of its local log sequences, mapping them into low-dimensional latent semantic vectors. Let the local log segment be x_i, and its representation be h_i; then the local encoding model can be expressed as:

$$h_i = f_\theta(x_i)$$

Here, f_θ represents the local parameterized representation network. To characterize the temporal and causal relationships in the log data, this method constructs time-dependent features on the local side and uses a sliding window mechanism to obtain a joint representation:

$$z_i = g(h_{i-W}, \ldots, h_i)$$

Where W is the window size and $g(\bullet)$ is the multi-scale aggregation function. This design ensures that each organization can extract system behavior semantics without exposing the source logs. This paper presents the model architecture diagram, as shown in Figure 1.

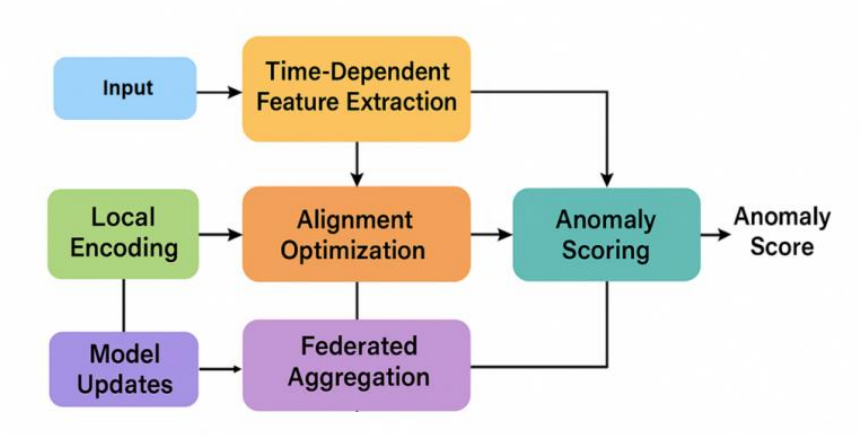

Figure 1. Overall model architecture diagram

To achieve cross-organizational collaboration capabilities, this method employs a federated update mechanism. Each participating organization uploads only model updates, not raw data, thereby ensuring privacy protection. The aggregation strategy uses parameter averaging; let θ_k be the local parameters of the k-th organization and θ be the globally aggregated parameters, then the process is as follows:

$$\theta = \frac{1}{K}\sum_{k=1}^{K}\theta_k$$

Here, K represents the number of organizations. Through repeated iterations, the representations of local models gradually converge towards global semantics, thereby improving cross-organizational anomaly understanding while maintaining data sovereignty and preventing data leakage.

To enhance cross-organizational representation consistency, this method introduces alignment optimization at the representation level. For semantic vectors z_a, z_b generated by different organizations, a consistency constraint is constructed to ensure that their distances converge in a shared subspace. This consistency loss is expressed as:

$$L_{align} = //z_a - z_b//_2^2$$

This constraint does not require sharing raw data, but rather shares semantics through covert channels, allowing collaborative modeling to still meet privacy and security requirements. At the same time, this approach reduces the bias caused by non-independent and identically distributed data, making the global representation more stable.

In terms of achieving anomaly sensitivity modeling, this method constructs a local anomaly scoring function, balancing both local and global semantics. The anomaly score is defined as:

$$s_i = //h_i - \mu//_2$$

Here, μ represents the semantic center, indicating the aggregated statistics of the normal mode. To prevent privacy leakage of the semantic center, this method employs a secure update mechanism during federated aggregation, calculating only the necessary statistics without sharing the original samples. The final training objective consists of a local self-supervised loss and a cross-organizational consistency loss:

$$L = L_{local} + \lambda L_{align}$$

Here, L_{local} represents the inter-organizational learning intensity coefficient. This objective allows for the construction of cross-domain representation capabilities without revealing the specific content of the logs, thus achieving privacy-preserving collaborative anomaly detection.

4. Experimental Results

4.1 Dataset

This study uses the HDFS (Hadoop Distributed File System) log anomaly detection dataset from LogHub as a unified open-source benchmark dataset. This dataset originates from the operational logs of a real distributed storage system and contains a large number of event templates generated over time, along with corresponding normal/abnormal labels. It is one of the most widely used and representative public datasets in the field of log anomaly detection, effectively reflecting the typical characteristics of "multi-component collaboration + temporal evolution + sparse anomalies" in cloud infrastructure, and is highly consistent with the research theme of "cloud log anomaly detection" in this paper.

From a data perspective, the HDFS dataset has a clear sequential structure, allowing logs to be aggregated into sample sequences based on block ID, session, or time window, facilitating the construction of the local encoding and time-dependent feature extraction processes in this method. Its anomaly patterns cover various types of faults or unexpected behaviors, including disk, network, and node interaction issues, supporting the need for anomaly sensitivity modeling based on low-dimensional semantic representations. It is also suitable for verifying the rationality of unsupervised/weakly supervised designs such as "semantic center + anomaly distance scoring."

Although this dataset itself comes from a single system environment, it is very suitable for constructing a cross-organizational federated learning simulation setting: the data can be divided into multiple "organizations/clients" based on different time periods, different sets of nodes, or different log template

subspaces, thus forming a non-independent and identically distributed multi-domain scenario. This maintains the advantages of open-source and reproducible data while remaining consistent with the method assumptions of "cross-organizational data remaining within their respective domains + federated representation alignment + collaborative anomaly detection," providing a standardized, scalable, and comparable experimental basis for subsequent research.

4.2 Experimental setup

All experiments are conducted on a single server equipped with 1 x NVIDIA A100 80GB GPU, 2 x Intel Xeon Gold CPUs (total 40 cores), and 256GB RAM. The operating system is Ubuntu 20.04 LTS. The implementation is based on Python 3.10, PyTorch 2.1.0, CUDA 12.1, and NCCL for efficient federated communication simulation. To ensure reproducibility, we fix the random seed to 42 and enable deterministic options in PyTorch where applicable.

For hyperparameters, the local encoder is a lightweight sequence model with hidden size 256 and dropout 0.1. The sliding window size is set to W=60, and the multi-scale aggregation uses three temporal scales (short/medium/long) implemented by pooled summaries over 5, 20, and 60 steps. We train with AdamW optimizer, learning rate 1e-3, weight decay 1e-4, batch size 128, and 50 local epochs per client. The federated setting uses K=5 clients, 100 global communication rounds, and full-client participation per round. The alignment loss weight is set to λ =0.5, and anomaly scoring uses an exponential moving average to update the semantic center with momentum 0.9.

4.3 Experimental Results

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

Table1: Comparative experimental results

Model	Acc	Precision	Recall	AUC
Fedas[53]	0.846	0.812	0.779	0.892
Fedbiot[54]	0.854	0.821	0.788	0.901
Openfedllm[55]	0.861	0.827	0.795	0.907
Fedllm-bench[56]	0.873	0.838	0.804	0.915
Fednlr[57]	0.882	0.846	0.811	0.921
SatFed[58]	0.889	0.854	0.820	0.927
Ours	0.907	0.874	0.842	0.944

Overall, the metrics show a steady upward trend as the model progresses from Fedas to SatFed, while Ours achieves the highest values in all four dimensions: Accuracy, Precision, Recall, and AUC. This indicates that our method possesses stronger comprehensive discriminative capabilities in cross-organizational log collaborative anomaly detection. Especially in a federated scenario, simultaneously improving accuracy and discriminative power without sacrificing recall often means that the model has not only learned a more stable "normal semantic center" but is also better at capturing shared abnormal deviation patterns across organizations. This aligns with the goal of "collaborative semantic representation under the premise of privacy protection" emphasized in this paper.

In terms of quantifiable gains, the improvement of Ours compared to the current strongest baseline SatFed is not dramatic but stable: Accuracy increased by approximately 0.018, Precision by approximately 0.020, Recall by approximately 0.022, and AUC by approximately 0.017. This "balanced improvement" better reflects the real challenges of cross-organizational cloud logs: inconsistent data distribution, differences in log templates, and sparse anomalies inhibit extreme leaps in any single metric. Therefore, this result is more

indicative of the systematic benefits brought about by representation alignment and temporal semantic aggregation, rather than an accidental advantage dependent on a specific threshold or local feature.

The improvement in Recall has more direct security implications. In cloud-native and cross-organizational operation and maintenance collaboration scenarios, false negatives are usually more damaging than false positives, as anomalies can propagate and amplify along the service dependency chain. Ours further increases Recall while maintaining a simultaneous improvement in Precision, suggesting that the model's characterization of the "semantic boundaries" of anomalies is clearer: it can both expand the coverage of abnormal patterns and avoid misclassifying more normal fluctuations as anomalies. This feature aligns with the method logic of "federated representation learning + representation consistency constraints," which strengthens the sensitivity to abnormal deviations across multiple organizations by sharing high-level semantics rather than sharing raw logs.

The leading AUC also further supports the robustness of the method under threshold-insensitive conditions. Due to the significant differences in alerting strategies and tolerance levels across different systems in a cross-organizational log environment, a model with a higher AUC is more likely to be safely deployed with different strategies on each organization's side. By simultaneously improving accuracy, precision, and recall, it can be reasonably inferred that this method effectively reduces representation drift caused by non-independent and identically distributed data while maintaining "privacy within the domain." This results in more stable global semantics and more reliable local anomaly scores, thus providing a more practically applicable collaborative foundation for privacy-preserving anomaly detection in cross-organizational cloud logs.

This paper also presents the results of hyperparameter sensitivity experiments. The experiments were conducted using the chosen optimizer, and the results are shown in Table 2.

Table 2: Results of hyperparameter sensitivity experiments

Optimizer	Acc	Precision	Recall	AUC
AdaGrad	0.872	0.836	0.802	0.918
SGD	0.883	0.847	0.811	0.923
Adam	0.895	0.857	0.823	0.933
AdamW	0.907	0.874	0.842	0.944

The sensitivity results indicate that different optimizers have a direct impact on the convergence stability and cross-organization semantic consistency of the federated representation learning framework. The overall trend shows gradual improvement from AdaGrad to SGD to Adam and AdamW. This suggests that in high-dimensional and sparse cloud log tasks with significant distribution differences, adaptive momentum and stronger regularization offer more stable representation extraction and collaborative alignment. This matches the training characteristics of the proposed method, which involves local encoding, temporal aggregation, and cross-organization consistency constraints.

The synchronous variation of the four metrics shows that AdamW achieves the best performance in Acc, Precision, Recall, and AUC. It provides a balanced optimization effect. For cross-organization log anomaly detection, accuracy alone is not sufficient to prove reliability. A more critical question is whether the model can continuously learn shared normal semantic structures while maintaining sensitivity to abnormal deviations, under the condition that privacy does not leave local domains. The advantage of AdamW implies that it suppresses local overfitting and prevents divergence of learning directions among clients. It improves the transferability of the semantic space after global aggregation.

The differences in Recall provide an application-oriented explanation. AdaGrad and SGD achieve relatively lower Recall. This indicates insufficient learning of rare anomaly patterns under federated settings. Anomalies in cloud native systems often propagate across services and show a temporal delay. If the optimizer cannot consistently enhance gradient contributions from weak signals, local representations become conservative. This dilutes anomaly semantics during global aggregation. The higher Recall of Adam and AdamW shows that they capture consistent abnormal deviation patterns across heterogeneous client data.

The stepwise increase in AUC further illustrates that the choice of optimizer influences model usability and robustness under different threshold settings. In cross-organization practice, participants may have different alert strategies and risk tolerance. Higher AUC means a more reliable ranking of anomaly scores and easier deployment for different organizations. Overall, AdamW as the default optimizer better supports the core objective of the federated representation learning framework. It forms stable, aligned, and anomaly-sensitive cross-organization semantic representations without sharing raw logs.

The representation dimension determines the semantic capacity that the model can handle and the separability of anomalous patterns, and is a key factor influencing cross-organizational consistency and local robustness in federated representation learning. A dimension that is too small may limit the ability to encode complex log semantics and temporal dependencies, while a dimension that is too large may lead to redundant representations and unstable training. Therefore, it is necessary to systematically examine how the sensitivity of the proposed algorithm to key metrics changes under different representation dimension settings. The experimental results are shown in Figure 2.

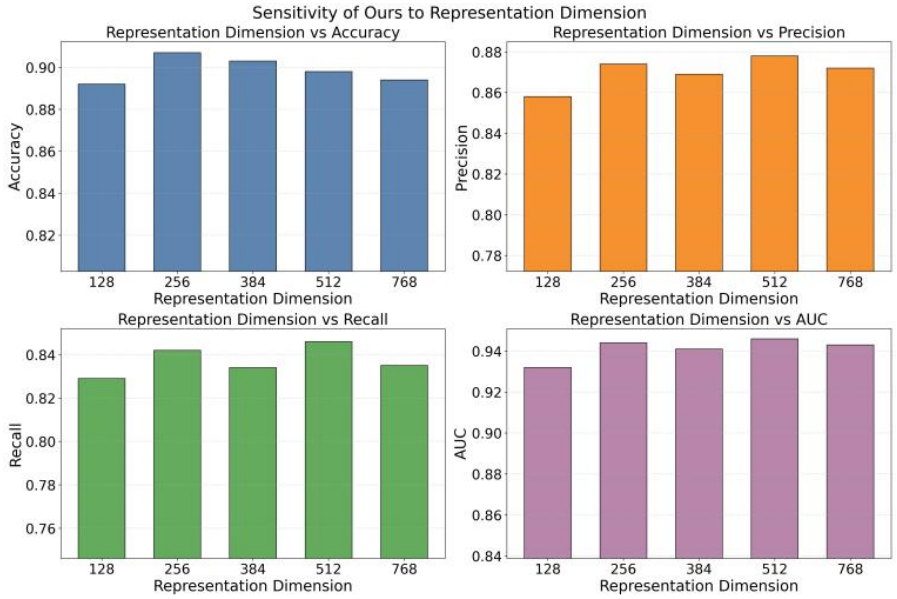

Figure 2. The influence of representation dimensions on experimental results

The figure shows that changes in representation dimensionality cause visible fluctuations across multiple metrics, but intermediate dimensionality produces more stable behavior. When the dimensionality increases from 128 to 256, all metrics improve notably. This indicates that moderately expanding semantic capacity helps the model encode cross-service semantics and temporal dependencies more effectively. It also provides a more reliable foundation for the shared semantic space in federated settings. This aligns with the intention to improve cross-organization anomaly representations without exposing raw logs.

For accuracy and precision, accuracy reaches a relatively optimal level near 256 and then decreases slightly as dimensionality continues to grow. This suggests that excessively high dimensionality may introduce redundancy. Local encodings on heterogeneous clients may drift slightly, reducing stability after global aggregation. Precision is better near 512. This implies that a richer representation space reduces false alarms

caused by normal fluctuations. It reflects a more fine-grained characterization of the normal semantic boundary.

The changes in recall and AUC further reinforce this interpretation. Recall reaches its peak near 512, showing more complete coverage of rare anomaly patterns. This benefits the detection of abnormal deviations that are shared across organizations. AUC also remains strong in the higher-dimensional interval. This indicates robust ranking of anomaly scores, which adapts better to different threshold strategies across organizations. This matches practical requirements in cloud native cross organization joint operation, where transferable alert capability is needed.

Across the four metrics, the range from 256 to 512 can be regarded as the more reasonable dimensionality interval. Lower dimensionality limits semantic capacity and weakens modeling of complex chain behaviors and temporal causal cues. Higher dimensionality improves local separability but increases instability caused by non-independent distributions and makes cross-client alignment more difficult. Therefore, setting dimensionality in an intermediate range and combining it with alignment optimization and federated aggregation better support cross-organization collaborative anomaly detection under privacy constraints.

Abnormal proportion changes directly alter the relative proportion of abnormal and normal patterns in the log sequence, thus affecting the ability of representation learning to capture rare outlier patterns. For federated representation learning, different organizations may have different prior distributions of anomalies, and this difference amplifies the alignment difficulties caused by non-independent and identically distributed data. Therefore, it is necessary to examine the sensitivity of the key metrics of our method under different anomaly proportion settings. The experimental results are shown in Figure 3.

The four subplots show that changes in the anomaly proportion produce interpretable effects on the decision behavior of the proposed method, while the model still maintains considerable stability. As the anomaly proportion increases from low to higher levels, accuracy shows a slight decline. This indicates that when anomaly samples occupy a larger share in the sequence, the statistical boundary of the normal semantic center is disturbed. The semantic distance-based decision then becomes more susceptible to boundary ambiguity. This behavior is consistent with real deployments, where anomaly distributions fluctuate with business phases and system load in cross-organization cloud log environments.

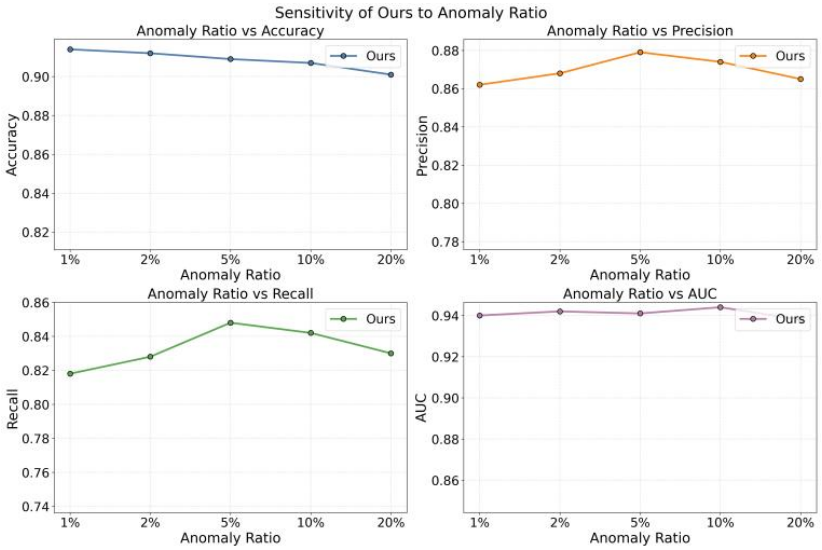

Figure 3. The impact of abnormal proportional changes on experimental results.

The four subplots show that changes in the anomaly proportion produce interpretable effects on the decision behavior of the proposed method, while the model still maintains considerable stability. As the anomaly proportion increases from low to higher levels, accuracy shows a slight decline. This indicates that when

anomaly samples occupy a larger share in the sequence, the statistical boundary of the normal semantic center is disturbed. The semantic distance-based decision then becomes more susceptible to boundary ambiguity. This behavior is consistent with real deployments, where anomaly distributions fluctuate with business phases and system load in cross-organization cloud log environments.

Precision first rises and then declines. It reaches a relatively high level when anomalies appear with moderate frequency. This means the model can more clearly differentiate semantic gaps between abnormal and normal signals under moderate anomaly strength, which reduces false alarms. When the anomaly proportion continues to increase, the decrease in precision suggests that anomaly patterns may become more diverse or semantically closer to some highly fluctuating normal patterns. The local representation thus loses the clearer boundaries observed at moderate levels. This also reflects that different prior anomaly characteristics across organizations can affect the purity of shared representations.

Recall shows a similar trend, performing better when the anomaly proportion is moderate. When anomalies are rare, the abnormal signals are sparse. The model needs stronger temporal and dependency cues to capture deviations reliably, so recall is limited. When the anomaly proportion reaches a moderate level, the semantic signals of anomalies become easier to reinforce through the encoder and temporal aggregation, leading to higher recall. The subsequent decline suggests that frequent anomalies weaken the contrast structure between abnormal and normal patterns. This can make it harder for alignment optimization to maintain consistent anomaly discrimination directions across multiple clients.

AUC remains high with small fluctuations across all proportions. This indicates robust ranking of anomaly scores and insensitivity to threshold selection. This is important for privacy-preserving cross-organization anomaly detection since different organizations often use different alert strategies and have different risk tolerances. A more stable AUC means that the shared representation can provide a consistent risk ranking. Taken together, the four metrics show that the method exhibits resilience to changes in anomaly proportion and performs more evenly at moderate levels. This supports the motivation of using federated representation learning to achieve stable anomaly characterization in cross-organization log collaboration.

Finally, this paper also presents a sensitivity analysis of the impact of log template noise intensity on accuracy (Acc), and the experimental results are shown in Figure 4.

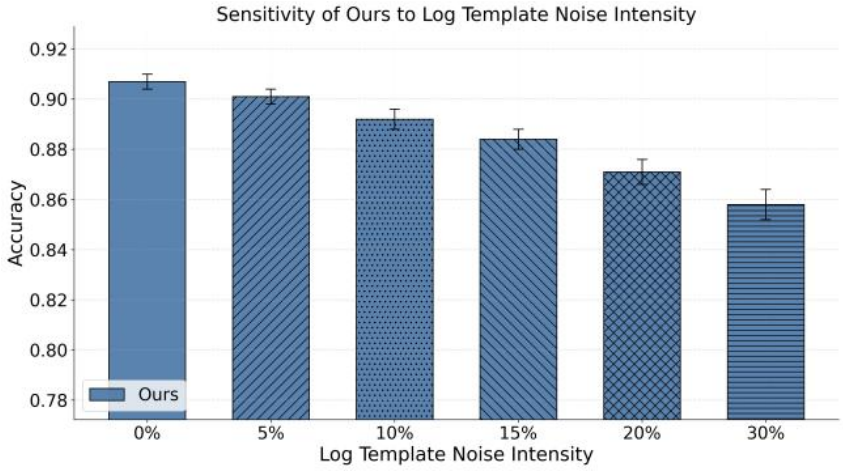

Figure 4. Sensitivity analysis of log template noise intensity on accuracy (Acc).

The figure shows that as the noise strength of log templates increases, the accuracy of the proposed method decreases steadily. This suggests that semantic disturbance at the template level weakens the model's ability to capture compact normal patterns. The method forms shared low-dimensional semantic representations through local encoding and temporal aggregation. Template noise disrupts semantic consistency and undermines the interpretability of event sequences. As a result, the estimation of the normal semantic center

shifts and the separability between normal and abnormal patterns decreases. This observation is consistent with practical situations in cloud logs, where template drift, version updates, and collection noise commonly occur.

From the perspective of cross-organization federated collaboration, the trend also shows that noise amplifies distribution differences among clients. This increases the difficulty of representation alignment. When log templates are already not perfectly consistent across organizations, additional noise makes semantic boundaries within the shared subspace more ambiguous. This affects the stability and generalization of global representations after aggregation. The result supports the importance of representation consistency emphasized in this study. It also suggests that in deployment, more robust template normalization, noise-aware enhancement, or adaptive alignment strategies could be combined to improve the reliability of privacy-preserving cross-organization log anomaly detection.

5. Conclusion

This study addresses the key challenge that cross-organization cloud logs cannot be collaboratively modeled under privacy constraints. It proposes a federated representation learning based anomaly detection framework. The goal is to share abnormal semantics and characterize risk without raw data, leaving local domains. The method builds structural log representations locally. It models temporal dependencies and applies sliding window aggregation to form low-dimensional latent semantics. Knowledge collaboration across organizations is achieved through federated updates. A consistency constraint at the representation level mitigates heterogeneous distribution and semantic drift. Each participant progressively converges toward a stable global semantic space without exposing raw logs. This enhances the distinguishability and transferability of abnormal patterns.

At the mechanism level, the framework shifts the core capability of anomaly detection from raw data aggregation to shared high-level semantic representations. It adopts a scoring logic based on semantic centers and deviation distances. It balances local robustness with global consistency. In cloud native and microservice environments, anomalies propagate, temporal patterns have multiple scales, and operational boundaries between organizations are clear. The design aligns naturally with coordination across teams, enterprises, or cloud platforms. It provides a technical foundation for sustained joint operation, joint risk control, and joint auditing. The work, therefore, has methodological significance and also offers practical value by advancing cloud security from isolated detection to collaborative intelligence.

In practical application scenarios, the study provides a deployable privacy-preserving collaboration paradigm. For multi-cloud and hybrid cloud operation, organizations can identify cross-domain abnormal coupling risks earlier through shared representations. For critical infrastructure and industrial cloud platforms, joint situational awareness can be established within compliance boundaries for supply chains and ecosystem partners. For regulatory technology systems, the framework supports cross-institution incident linkage and evidence chain analysis with stronger semantic reliability. As cloud services scale and cross-organization collaboration becomes routine, sharing representations rather than sharing raw data has the potential to become a core component of privacy-friendly cloud security and intelligent operation.

Future work may progress in several dimensions. One direction is incorporating finer structural priors and causal modeling to improve the semantic explainability of cross-service propagation chains. This enhances abnormal localization and root cause analysis. Another direction is adding stronger privacy protection and communication compression strategies. This reduces cross-organization training cost while maintaining security, improving deployment feasibility in real environments. A third direction is extending to multi-source observation settings by integrating metrics, logs, and trace signals. This improves early detection of weak anomalies and emerging risks. Overall, the study provides a clear technical route and extensible framework for privacy-preserving cross-organization anomaly detection on cloud logs. It has the potential to generate lasting impact in cloud native security, intelligent operation, and collaborative governance across ecosystems.

References

[1] B. Li, S. Ma, R. Deng, et al., "Federated Anomaly Detection on System Logs for the Internet of Things: A Customizable and Communication-Efficient Approach," IEEE Transactions on Network and Service Management, vol. 19, no. 2, pp. 1705-1716, 2022.

[2] T. H. Shin and S. H. Kim, "Utility Analysis About Log Data Anomaly Detection Based on Federated Learning," Applied Sciences, vol. 13, no. 7, p. 4495, 2023.

[3] R. Xu, H. Miao, S. Wang, et al., "PeFAD: A Parameter-Efficient Federated Framework for Time Series Anomaly Detection," Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pp. 3621-3632, 2024.

[4] 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.

[5] W. Meng, Y. Liu, Y. Zhu, S. Zhang, D. Pei, Y. Liu, et al., "Loganomaly: Unsupervised Detection of Sequential and Quantitative Anomalies in Unstructured Logs," IJCAI, vol. 19, no. 7, pp. 4739-4745, 2019.

[6] X. Zhang, Y. Xu, Q. Lin, B. Qiao, H. Zhang, Y. Dang, et al., "Robust Log-Based Anomaly Detection on Unstable Log Data," Proceedings of the 2019 27th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, pp. 807-817, 2019.

[7] S. He, J. Zhu, P. He and M. R. Lyu, "Experience Report: System Log Analysis for Anomaly Detection," 2016 IEEE 27th International Symposium on Software Reliability Engineering (ISSRE), pp. 207-218, 2016.

[8] V. H. Le and H. Zhang, "Log-Based Anomaly Detection Without Log Parsing," 2021 36th IEEE/ACM International Conference on Automated Software Engineering (ASE), pp. 492-504, 2021.

[9] P. He, J. Zhu, Z. Zheng and M. R. Lyu, "Drain: An Online Log Parsing Approach With Fixed Depth Tree," 2017 IEEE International Conference on Web Services (ICWS), pp. 33-40, 2017.

[10] J. Zhu, S. He, J. Liu, P. He, Q. Xie, Z. Zheng and M. R. Lyu, "Tools and Benchmarks for Automated Log Parsing," 2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP), pp. 121-130, 2019.

[11] K. 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 & Data Mining, pp. 387-395, 2018.

[12] P. Malhotra, L. Vig, G. Shroff and P. Agarwal, "Long Short Term Memory Networks for Anomaly Detection in Time Series," Proceedings, vol. 89, no. 9, p. 94, 2015.

[13] Y. Bengio, A. Courville and P. Vincent, "Representation Learning: A Review and New Perspectives," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 8, pp. 1798-1828, 2013.

[14] L. Ruff, R. Vandermeulen, N. Goernitz, L. Deecke, S. A. Siddiqui, A. Binder, et al., "Deep One-Class Classification," International Conference on Machine Learning, pp. 4393-4402, 2018.

[15] R. Chalapathy and S. Chawla, "Deep Learning for Anomaly Detection: A Survey," arXiv preprint arXiv:1901.03407, 2019.

[16] M. Landauer, S. Onder, F. Skopik, et al., "Deep Learning for Anomaly Detection in Log Data: A Survey," Machine Learning with Applications, vol. 12, p. 100470, 2023.

[17] X. Wei, J. Wang, C. Sun, et al., "Log-Based Anomaly Detection for Distributed Systems: State of the Art, Industry Experience, and Open Issues," Journal of Software: Evolution and Process, vol. 36, no. 8, p. e2650, 2024.

[18] R. Al-Amri, R. K. Murugesan, M. Man, A. F. Abdulateef, M. A. Al-Sharafi and A. A. Alkahtani, "A Review of Machine Learning and Deep Learning Techniques for Anomaly Detection in IoT Data," Applied Sciences, vol. 11, no. 12, p. 5320, 2021.

[19] B. McMahan, E. Moore, D. Ramage, S. Hampson and B. A. y Arcas, "Communication-Efficient Learning of Deep Networks From Decentralized Data," Artificial Intelligence and Statistics, pp. 1273-1282, 2017.

[20] P. Kairouz and H. B. McMahan, "Advances and Open Problems in Federated Learning," Foundations and Trends in Machine Learning, vol. 14, no. 1-2, pp. 1-210, 2021.

[21] Y. Liu, S. Garg, J. Nie, Y. Zhang, Z. Xiong, J. Kang and M. S. Hossain, "Deep Anomaly Detection for Time-Series Data in Industrial IoT: A Communication-Efficient On-Device Federated Learning Approach," IEEE Internet of Things Journal, vol. 8, no. 8, pp. 6348-6358, 2020.

[22] T. D. Nguyen, S. Marchal, M. Miettinen, et al., "DÏoT: A Federated Self-Learning Anomaly Detection System for IoT," 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), pp. 756-767, 2019.

[23] J. Cai, Y. Zhang, J. Fan, et al., "Lg-FGAD: An Effective Federated Graph Anomaly Detection Framework," Proceedings of the International Joint Conference on Artificial Intelligence, pp. 3760-3769, 2024.

[24] D. Wu, "Federated Deep Learning With Contrastive Representation for Node State Identification in Distributed Systems," Transactions on Computational and Scientific Methods, vol. 4, no. 8, 2024.

[25] T. Chen, S. Kornblith, M. Norouzi and G. Hinton, "A Simple Framework for Contrastive Learning of Visual Representations," International Conference on Machine Learning, pp. 1597-1607, 2020.

[26] B. Barlocker and X. Yan, "Contrastive Representation Learning for Anomaly Detection in Cloud-Based Backend Services," Artificial Intelligence and Computing Innovations, vol. 1, no. 2, 2021.

[27] T. N. Kipf and M. Welling, "Semi-Supervised Classification With Graph Convolutional Networks," arXiv preprint arXiv:1609.02907, 2016.

[28] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio and Y. Bengio, "Graph Attention Networks," arXiv preprint arXiv:1710.10903, 2017.

[29] J. Dean and L. A. Barroso, "The Tail at Scale," Communications of the ACM, vol. 56, no. 2, pp. 74-80, 2013.

[30] J. Von Kistowski, S. Eismann, N. Schmitt, A. Bauer, J. Grohmann and S. Kounev, "Teastore: A micro-service reference application for benchmarking, modeling and resource management research," Proceedings of the 2018 IEEE 26th International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), pp. 223-236, 2018.

[31] F. Chen, "AI-Augmented Anomaly Detection via Generative Distribution Modeling and Uncertainty Quantification in Cloud Systems," 2024.

[32] Z. Qiu, "A Multi-Scale Deep Learning and Uncertainty Estimation Framework for Comprehensive Anomaly Detection in Cloud Environments," 2023.

[33] Y. Ma, "Anomaly Detection in Microservice Environments via Conditional Multiscale GANs and Adaptive Temporal Autoencoders," Transactions on Computational and Scientific Methods, vol. 4, no. 10, 2024.

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

[35] Z. Qiu, "Time Series and Graph Structure Fusion for AI-Based Anomaly Detection in Microservice Environments," Journal of Computer Technology and Software, vol. 3, no. 7, 2024.

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

[37] Y. Deng, "Transfer Methods for Large Language Models in Low-Resource Text Generation Tasks," 2024.

[38] H. Liu, "Structural Regularization and Bias Mitigation in Low-Rank Fine-Tuning of LLMs," Transactions on Computational and Scientific Methods, vol. 3, no. 2, 2023.

[39] Y. Li, "Task-Aware Differential Privacy and Modular Structural Perturbation for Secure Fine-Tuning of Large Language Models,"2024.

[40] A. Xie, "Adaptive Privacy-Aware Federated Language Modeling for Collaborative Electronic Medical Record Analysis," 2024.

[41] Y. Wang, "Integrating Large Language Models and Knowledge Graphs for Intelligent Financial Regulatory Risk Identification," Transactions on Computational and Scientific Methods, vol. 4, no. 11, 2024.

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

[43] Y. Wang, "Semantic-Driven Large Model Scheduling for Distributed Systems via Unified Representation and Policy Generation," 2024.

[44] Q. Gan, "Large Language Model Framework for Multi-Document Financial Anomaly Detection in Intelligent Auditing via Semantic Mapping and Risk Reasoning," 2024.

[45] X. Sun, Y. Yao, X. Wang, P. Li and X. Li, "AI-Driven Health Monitoring of Distributed Computing Architecture: Insights From XGBoost and SHAP," 2024 4th International Conference on Communication Technology and Information Technology (ICCTIT), pp. 480-484, 2024.

[46] Y. Xing, "Enhancing Advertising Recommendation Performance via Integrated Causal Inference and Exposure Bias Correction," 2023.

[47] M. Wang, "Multi-Level Attention and Sequence Modeling for Dynamic User Interest Representation in Real-Time Advertising Recommendation," Transactions on Computational and Scientific Methods, vol. 3, no. 2, 2023.

[48] S. Pan, T. Hu, S. Sun, J. Yuan and J. Luo, "Help Oneself in Helping the Others: The Ecology of Online Support Groups," 2019 IEEE International Conference on Big Data (Big Data), pp. 2418-2427, 2019.

[49] A. M. Jones, G. Sahin, Z. W. Murdock, Y. Ge, A. Xu, Y. Li, et al., "USC-DCT: A Collection of Diverse Classification Tasks," Data, vol. 8, no. 10, p. 153, 2023.

[50] Y. Hu, "Autonomous Agent Architecture for Complex Tasks via Hierarchical Planning and Language Model Reasoning," 2024.

[51] J. Lai, "Attention Alignment under Logical Constraints for Reliable Financial Statement Reasoning," 2024.

[52] Y. Huang, "Explainable Cognitive Multi-Agent AI for Joint Intention Modeling in Complex Task Planning," 2024.

[53] X. Yang, W. Huang and M. Ye, "Fedas: Bridging Inconsistency in Personalized Federated Learning," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11986-11995, 2024.

[54] F. Wu, Z. Li, Y. Li, et al., "Fedbiot: LLM Local Fine-Tuning in Federated Learning Without Full Model," Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pp. 3345-3355, 2024.

[55] R. Ye, W. Wang, J. Chai, et al., "Openfedllm: Training Large Language Models on Decentralized Private Data via Federated Learning," Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pp. 6137-6147, 2024.

[56] R. Ye, R. Ge, X. Zhu, et al., "Fedllm-Bench: Realistic Benchmarks for Federated Learning of Large Language Models," Advances in Neural Information Processing Systems, vol. 37, pp. 111106-111130, 2024.

[57] H. Wang, P. Zheng, X. Han, et al., "Fednlr: Federated Learning With Neuron-Wise Learning Rates," Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pp. 3069-3080, 2024.

[58] E. Diao, J. Ding and V. Tarokh, "Heterofl: Computation and communication efficient federated learning for heterogeneous clients," arXiv preprint arXiv:2010.01264, 2020.