xHybrid: Explainable Hybrid Anomaly Detection in Microservice Traces via Grammar-Neural Fusion

Tanvir Rahman (22241134)
Department of Computer Science and Engineering
Brac University, Dhaka, Bangladesh
tanvir.rahman1@g.bracu.ac.bd

Abstract

Microservice architectures generate millions of distributed traces daily, requiring effective anomaly detection for system reliability. Existing approaches face a fundamental accuracy-interpretability trade-off: deep learning methods achieve high accuracy but lack explainability, while rule-based methods are interpretable but sacrifice performance. I propose **xHybrid**, a hybrid framework combining grammar-based pattern extraction, graph neural networks, and temporal modeling with multi-level explainability. The approach integrates SAX-Sequitur grammar for interpretable rules, Graph Attention Networks for service dependencies, LSTM for temporal patterns, and attention-based fusion with SHAP explanations. Preliminary evaluation on 10,000 Train-Ticket traces achieves 99.2% accuracy and F1-score of 0.979, with explainability through grammar rules, attention weights, and SHAP feature importance.

Keywords: Anomaly Detection, Microservices, Explainable AI, Graph Neural Networks, Distributed Traces

1 Introduction and Motivation

Microservice architectures are the standard for cloud-native applications, generating millions of distributed traces daily. Anomaly detection is critical for system reliability, but existing approaches face a fundamental **accuracy vs. interpretability trade-off**:

- **Deep Learning** (Twin Graph [6], DeepTraLog [7]): High accuracy, but black-box
- **Rule-Based** (D'Angelo [4]): Interpretable, but lower accuracy
- **Recent Methods** (FC-ADL [1]): Fast, but limited explainability

Research Gap: No existing work combines structural (GNN), temporal (LSTM), and syntactic (grammar) features with comprehensive multi-level explainability.

2 Literature Review

Table 1 summarizes recent advances in microservice anomaly detection.

3 Proposed Approach: xHybrid

I propose **xHybrid**, a hybrid framework combining:

1. **Grammar Branch** (SAX + Sequitur) – Interpretable syntactic patterns
2. **GNN Branch** (Graph Attention) – Service dependency modeling
3. **LSTM Branch** (Bidirectional) – Temporal sequence patterns
4. **Attention Fusion** – Learnable component weights
5. **Multi-level xAI** – Grammar rules + Attention + SHAP

3.1 Architecture

Figure 1 illustrates the xHybrid architecture with three parallel branches fused via attention mechanism.

Table 1: State-of-the-Art Comparison in Microservice Anomaly Detection

Paper	Venue	Method	Key Contribution	Best Metric	xAI
FC-ADL [1]	SoCC 2025	Functional connectivity	Causal discovery, 152× faster	F1: 0.95	Partial
Few-Shot [2]	arXiv 2024	Meta-learning	Cross-system transfer	Acc: 93.26%	None
CHASE [3]	FSE 2024	Causal hypergraph	Root cause analysis, +36.2% A@1	–	Partial
D'Angelo [4]	ICPE 2024	SAX + Sequitur	Grammar-based patterns	–	Full
ServiceAnomaly [5]	JSS 2024	Traces + metrics	Multi-modal DAG	F1: 0.86	None
MSTGAD [6]	ASE 2023	Multi-modal GNN	Twin graph learning	F1: 0.961	Partial
DeepTraLog [7]	ICSE 2022	Trace-log GNN	Graph-based SVDD	F1: 0.954	None
TraceAnomaly [8]	ISSRE 2020	Deep Bayesian	Service-level VAE	P/R: 0.97	None

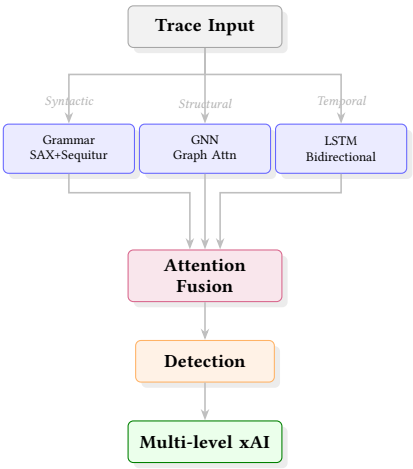

Figure 1: xHybrid: Three-branch fusion with multi-level explainability

3.2 Explainability Stack

- **Level 1:** Grammar rules – Human-readable patterns
- **Level 2:** Attention weights – Service importance
- **Level 3:** SHAP values – Feature contributions

3.3 Datasets

- **Train-Ticket:** Benchmark microservice system with 41 services, 22 replicated fault scenarios
- **Fault Types:** cpu, mem, disk, delay, loss, socket (6 types)
- **Cross-system:** Sock-Shop (13 services, different architecture)
- **Current Evaluation:** 10K sample from processed data

4 Preliminary Results

Evaluation: 10,000 Train-Ticket traces (7,000 train, 1,500 val, 1,500 test)
Note: Results are on a sample dataset for proof-of-concept. Full dataset evaluation is planned as future work.

5 Limitations

1. **Sample-based evaluation** – Current results are on 10K traces; full dataset evaluation is needed for conclusive comparison
2. **Single-dataset** – Cross-system validation on Sock-Shop is planned
3. **Trace-only input** – Future work includes multi-modal fusion (traces + logs + metrics)

Table 2: Performance Comparison (xHybrid results are preliminary on 10K sample)

Method	Accuracy	Precision	Recall	F1	xAI	Notes
xHybrid*	**99.2%**	**0.979**	**0.979**	**0.979**	**Full**	Preliminary, 10K samples, 38s training
FC-ADL	–	–	–	0.95	Partial	152× faster inference
Twin Graph	–	–	–	0.961	Partial	Multi-modal GNN
Few-Shot	93.26%	–	–	–	None	Cross-system transfer

*Direct comparison requires evaluation on same dataset

6 Future Work

1. Full-scale evaluation on complete Train-Ticket dataset
2. Ablation study to quantify each component's contribution
3. Cross-system validation on Sock-Shop, DeathStarBench
4. xAI visualizations (attention heatmaps, SHAP plots)
5. Per-fault-type analysis (cpu, mem, disk, delay, loss, socket)
6. Real-time deployment and latency evaluation

7 Conclusion

xHybrid addresses the gap between accuracy and interpretability in microservice anomaly detection. The explainability focus (grammar + attention + SHAP) differentiates it from recent methods. Preliminary results on a 10K sample show promising accuracy with multi-level explainability, though full-scale evaluation is needed for conclusive comparison with state-of-the-art methods.

References

[1] G. Winchester, G. Parisis, and L. Berthouze, "FC-ADL: Efficient Microservice Anomaly Detection and Localisation Through Functional Connectivity," in *Proc. ACM Symp. Cloud Computing (SoCC)*, 2025.

[2] Y. Wang, M. V. Mäntylä, S. Demeyer, M. Beyazit, J. Kisaakye, and J. Nyssölä, "Cross-System Categorization of Abnormal Traces in Microservice-Based Systems via Meta-Learning," *arXiv:2403.18998*, 2024.

[3] Z. Zhao, Z. Wang, T. Zhang, et al., "CHASE: A Causal Hypergraph based Framework for Root Cause Analysis in Multimodal Microservice Systems," in *Proc. ACM SIGSOFT Int. Symp. Foundations of Software Engineering (FSE)*, 2024.

[4] A. D'Angelo and A. Di Marco, "Grammar-Based Anomaly Detection of Microservice Systems Execution," in *Proc. ACM/SPEC Int. Conf. Performance Engineering (ICPE)*, 2024.

[5] M. Panahandeh, A. Hamou-Lhadj, M. Hamdaqa, and J. Miller, "ServiceAnomaly: An anomaly detection approach in microservices using distributed traces and profiling metrics," *J. Systems and Software*, vol. 209, 2024.

[6] J. Huang, Y. Yang, H. Yu, J. Li, and X. Zheng, "Twin Graph-based Anomaly Detection via Attentive Multi-Modal Learning for Microservice System," in *Proc. IEEE/ACM Int. Conf. Automated Software Engineering (ASE)*, 2023.

[7] C. Zhang, X. Peng, C. Sha, et al., "DeepTraLog: Trace-Log Combined Microservice Anomaly Detection through Graph-based Deep Learning," in *Proc. IEEE/ACM Int. Conf. Software Engineering (ICSE)*, 2022.

[8] P. Liu, H. Xu, Q. Ouyang, et al., "Unsupervised Detection of Microservice Trace Anomalies through Service-Level Deep Bayesian Networks," in *Proc. IEEE Int. Symp. Software Reliability Engineering (ISSRE)*, 2020.