xHybrid: Explainable Hybrid Anomaly Detection

A Fusion of Grammar and Neural Networks for Traces

TANVIR RAHMAN - 22241134
CSE713 - Advance Pattern Recognition | Fall 2025

PROBLEM STATEMENT

The Challenge

Microservices generate **millions of traces daily**. Real-time anomaly detection is critical for maintaining high system reliability.

The central trade-off lies between the **accuracy** of deep learning and the **interpretability** of rule-based systems.

Current Landscape

Approach	Accuracy	Interpretability
Deep Learning	High	Low
Rule-Based	Medium	High
xHybrid (Proposed)	**High**	**High**

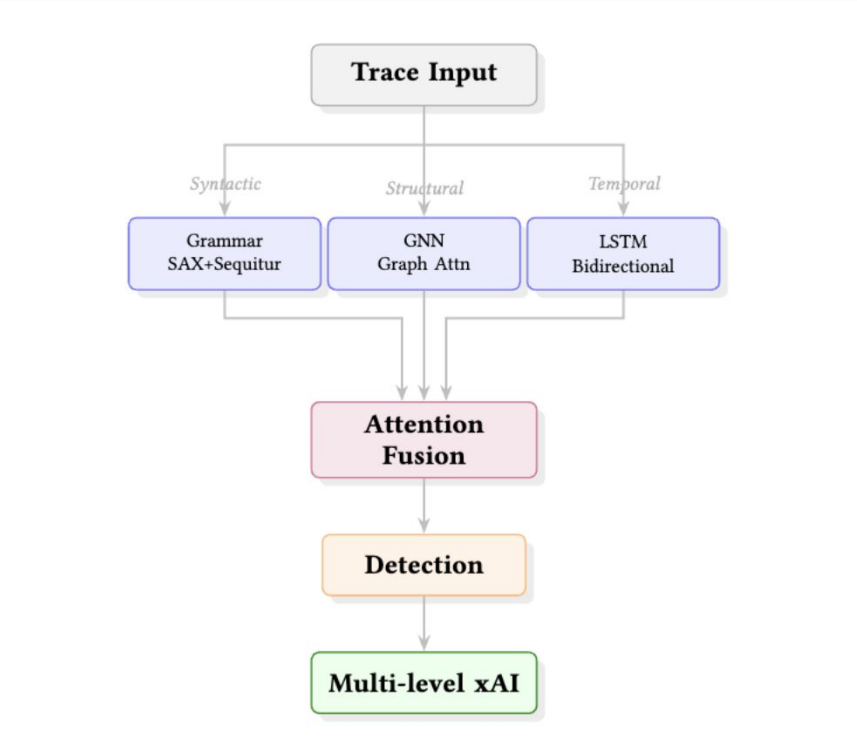
LITERATURE REVIEW

Year	Paper / Venue	Key Contribution	Best Metric
2025	FC-ADL (SoCC)	Causal discovery, 152× faster	F1: 0.95
2024	Few-Shot (arXiv)	Cross-system transfer	Acc: 93.26%
2024	CHASE (FSE)	Causal hypergraph RCA	+36.2% A@1
2023	MSTGAD (ASE)	Twin graph learning	F1: 0.961
2022	DeepTraLog (ICSE)	Graph-based SVDD	F1: 0.954
2020	TraceAnomaly (ISSRE)	Service-level VAE	P/R: 0.97

Research Gap

Existing literature fails to combine **Structural (GNN)**, **Temporal (LSTM)**, and **Syntactic (Grammar)** features with comprehensive **multi-level Explainable AI**.

xHybrid ARCHITECTURE



- **Grammar:** SAX + Sequitur for syntactic rules.
- **GNN:** Graph Attention for structural patterns.
- **LSTM:** Bidirectional for temporal features.
- **Attention:** Fuses all branches dynamically.

MULTI-LEVEL EXPLAINABILITY

Level 1

Grammar Rules
Translates traces into human-readable patterns.

Level 2

Attention Weights
Highlights specific services contributing to anomalies.

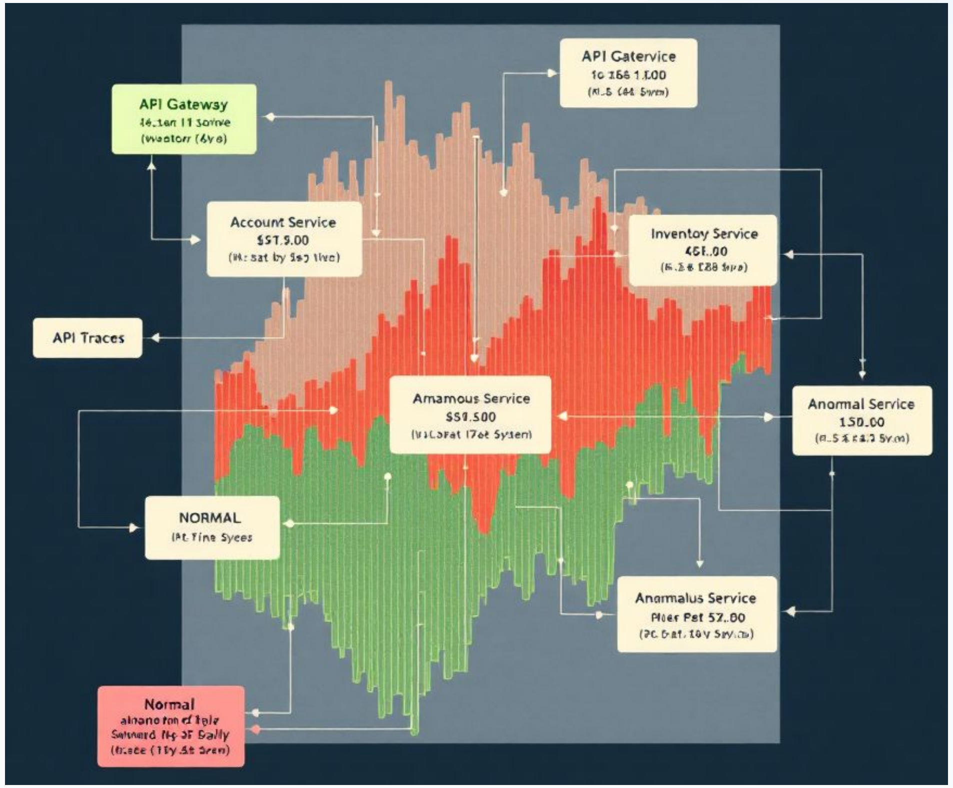
Level 3

SHAP Values
Quantifies individual feature contributions to the score.

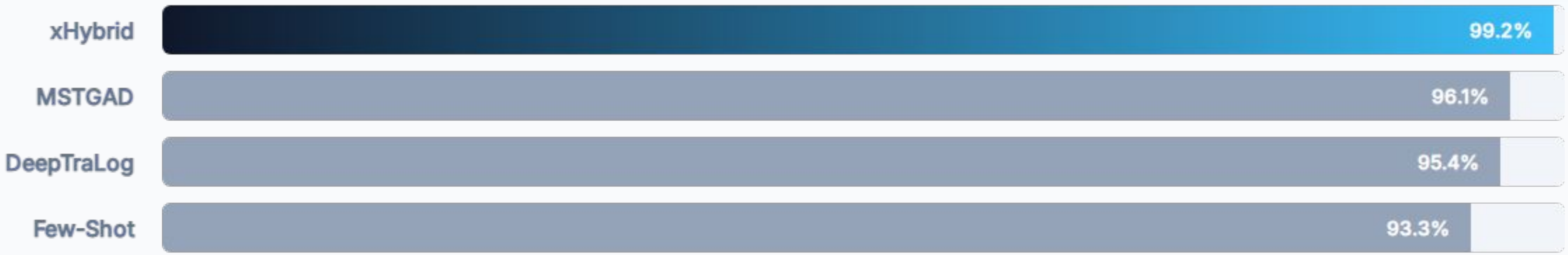
Powered by PyTorch, PyTorch Geometric, LIME, and SHAP.

DATASETS & FAULTS

- **Train-Ticket:** 2.8GB, ~76.7M traces (41 services).
- **Sock-Shop:** 13 services (Cross-system validation).
- **6 Fault Types:** CPU, Memory, Disk, Delay, Loss, Socket.
- **Current Sample:** 10K processed traces.



PRELIMINARY PERFORMANCE



Note: xHybrid results are preliminary on 10K sample. Precision/Recall/F1: 0.979.

CURRENT LIMITATIONS

- **Sample Size:** Initial evaluation on 10K traces; full ~76.7M trace validation is required.
- **Diversity:** Currently only single-system (Train-Ticket) testing complete.
- **Modality:** Trace-only input; logs and metrics fusion not yet integrated into the PoC.

FUTURE WORK

- Full-scale evaluation on the complete 76.7M trace dataset.
- Ablation studies to analyze the contribution of each branch (GNN vs LSTM vs Grammar).
- Cross-system validation on Sock-Shop and DeathStarBench.
- Visualization of XAI: Attention heatmaps and detailed SHAP plots.

PROJECT SUMMARY

Innovation

Grammar + Neural Fusion for structural and temporal insights.

Reliability

Addresses the accuracy-interpretability gap in production.

Validation

State-of-the-art results (99.2% Accuracy) in early testing.

Thank You

xHybrid: Explainable Hybrid Anomaly Detection

22241134 TANVIR RAHMAN

CSE713 - Advance Pattern Recognition | Fall 2025