ASSESSING ML-BASED ANOMALY DETECTION ACROSS INDIVIDUAL TELEMETRY MODALITIES IN CLOUD-NATIVE B5G NETWORKS

STOYAN KUTSAROV

RESPONSIBLE PROFESSOR
NITINDER MOHAN

INTRODUCTION

Cloud-native systems generate large volumes of telemetry data, including logs, metrics, and traces, which are essential for monitoring and diagnosing failures in distributed environments. Logs record discrete events and execution paths; metrics capture continuous system state; distributed traces expose inter-service latency and error propagation. Existing anomaly detection research predominantly addresses each modality in isolation and is done on datasets that don't reflect the structure of a 5G core. This project investigates how effectively machine learning models can detect anomalies using each telemetry modality independently, with the goal of identifying their strengths, limitations, and suitability for different types of system faults.

SUBQUESTIONS

- How effectively can ML models detect anomalies from logs/metrics/ traces when each modality is analyzed independently?
- How does anomaly detection performance vary across modalities for different types of faults or abnormal behavior?
- Which modalities are most informative for identifying specific types of anomalies?
- How robust are anomaly detection models to noise and variability in each modality

METHODOLOGY

Dataset
- Open5GS and UERANSIM (10 simulated UEs)
- Deployed in kind on a single host.
- 600s baseline, 300s fault, 300s recovery
- 22 faults

Telemetry
- Logs (Loki)
- Metrics (Prometheus)
- Traces (Jaeger)

Labeling
- Anomalous if collected during fault
- Normal otherwise

Log Models
- LogBERT
- Logs2Graphs
- DeepLog
- LogRobust
- Isolation Forest

Metric Models
- PCA
- USAD
- OmniAnomaly
- TranAD
- AnomalyTransformer

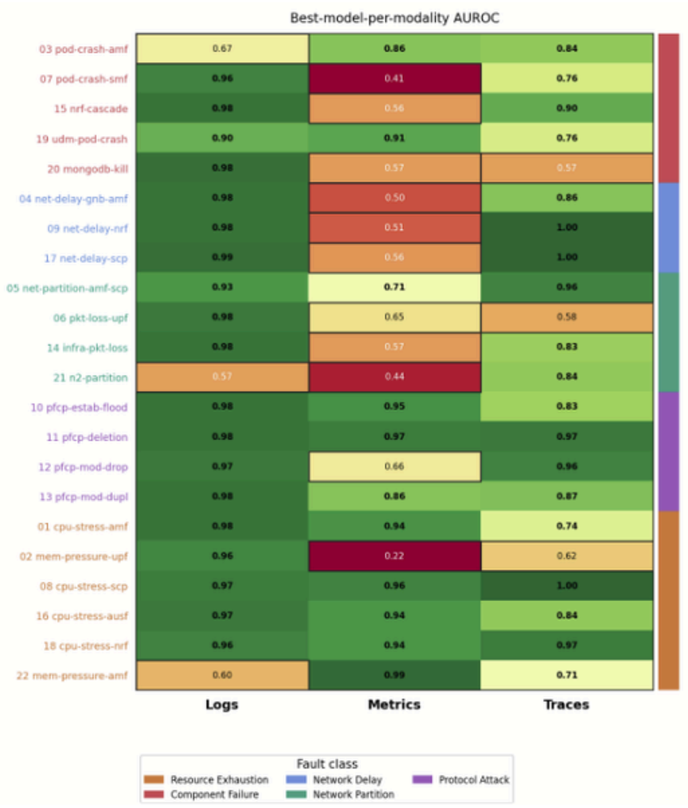
Traces Models
- TraceRPCA
- TraceAnomaly
- GAL-MAD
- TraceDAE
- TraceSieve

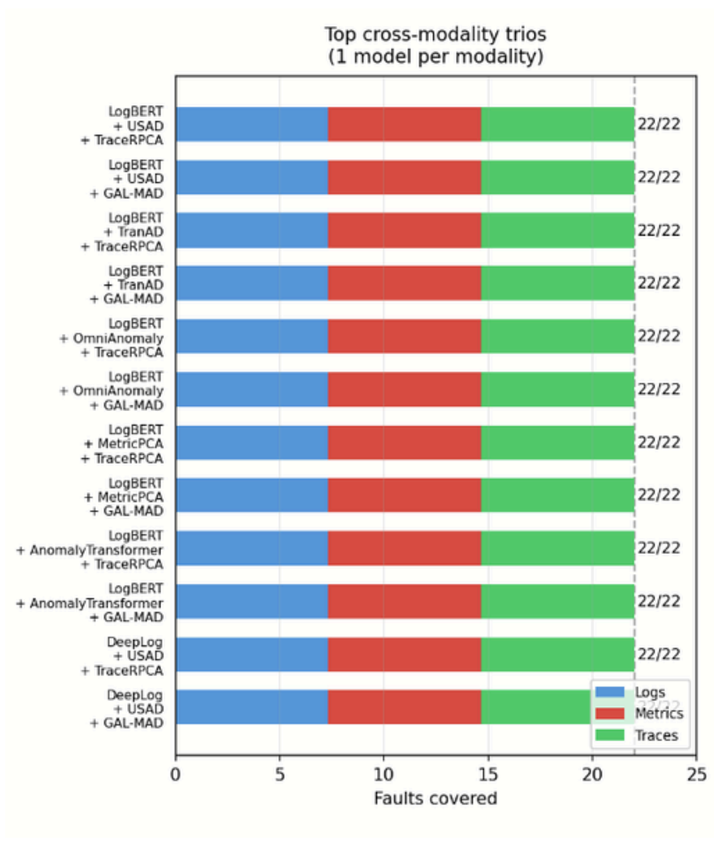
EXPERIMENTS

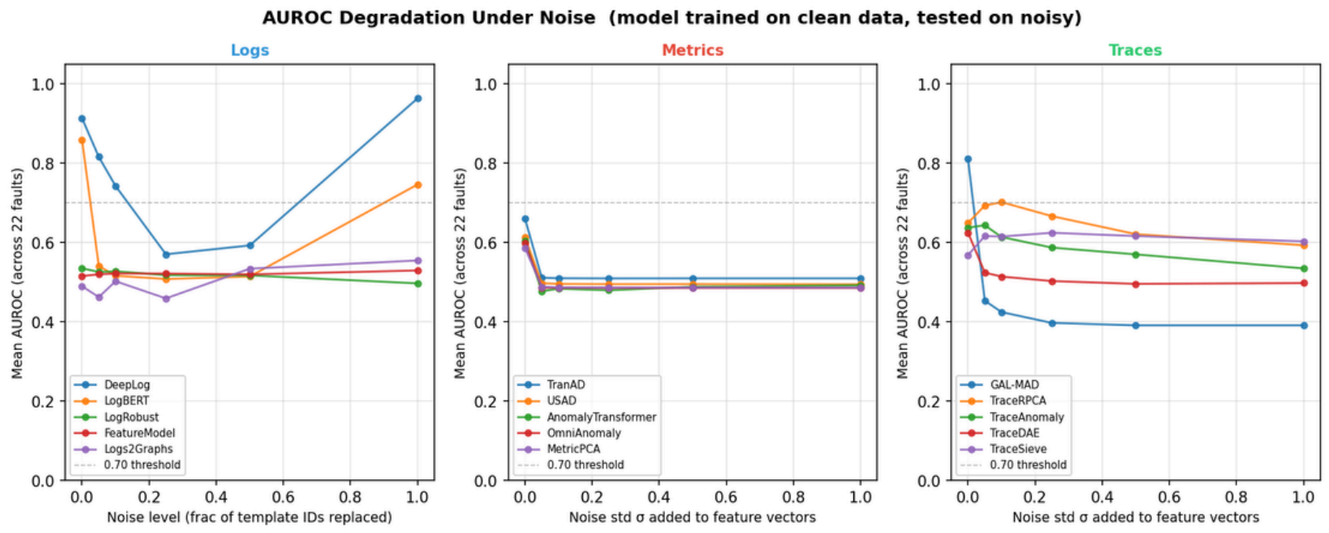
- Within each modality, multiple models spanning statistical baselines through deep generative architectures are trained unsupervised

on pre-fault normal data and evaluated per fault

- All models share identical training/test splits and evaluation metrics within a modality, enabling direct comparison
- Results are reported per fault as heatmaps of AUROC, Average Precision and Recall







DISCUSSION & CONCLUSION

Key Findings
- Log models dominate: DeepLog detects 19/22 faults. Three models sufficient to capture all faults.
- Metric models hampered by cluster-level aggregation, dropping 5G control-plane features does not degrade performance, CPU usage metrics very important for them
- Trace models cluster at 0.60 - 0.66 AUROC, TraceSieve most robust to feature dropout
- Good models are not resilient under noise, performance degrades rapidly and may lead to false positives for logs.

Limitations
- Dataset limited at present, could be made robust in more ways including trials, types of faults, number of UEs
- Phase-level labelling introduces noise. Records within the fault window that are behaviourally normal are incorrectly marked as anomalous
- Metric data gathered in a way that's not as useful as it could be for the task at hand.

Lin, Q. et al. "Log Clustering Based Problem Identification for Online Service Systems." ICSE Companion, 2016.
Du, M. et al. "DeepLog: Anomaly Detection and Diagnosis from System Logs through Deep Learning." ACM CCS, 2017.
Meng, W. et al. "LogAnomaly: Unsupervised Detection of Sequential and Quantitative Anomalies in Unstructured Logs." IJCAI, 2019.
Su, Y. et al. "Robust Anomaly Detection for Multivariate Time Series through Stochastic Recurrent Neural Network." ACM KDD, 2019.
Jin, M. et al. "An Anomaly Detection Algorithm for Microservice Architecture Based on Robust Principal Component Analysis." IEEE Access, 2020.
Liu, P. et al. "Unsupervised Detection of Microservice Trace Anomalies through Service-Level Deep Bayesian Networks." IEEE ISSRE, 2020.
Li, Z. et al. "Practical Root Cause Localization for Microservice Systems via Trace Analysis." IEEE/ACM IWQoS, 2021.
Yu, G. et al. "MicroRank: End-to-End Latency Issue Localization with Extended Spectrum Analysis in Microservice Environments." ACM WWW, 2021.