INTERNATIONAL JOURNAL OF COMPUTER SCIENCE & INFORMATION SYSTEM

Volume10 Issue10, October-2025, pg. 117-123

Published Date: - 31-10-2025

E-ISSN: 2536-7919
P-ISSN: 2536-7900
SJIF 2019: 4.58 2020: 5.046 2021: 5.328

Intelligent Log-Driven Anomaly Detection and Failure Prediction in Cloud-Native Microservices for 6G-Enabled Ultra-Low-Latency Systems

Dr. Adrian Keller

Department of Computer Science, University of Zurich, Switzerland

Abstract: The rapid convergence of cloud-native microservices, edge computing, and emerging 6G network architectures has intensified the need for intelligent, real-time anomaly detection and failure prediction mechanisms. As distributed systems evolve toward ultra-reliable and low-latency communication paradigms, particularly in support of holographic communications and immersive applications, traditional monitoring approaches become insufficient. This study presents a comprehensive research framework integrating log-based modeling, machine learning-assisted service boundary detection, semi-Markov failure prediction, correlation-driven anomaly analysis, and edge-enabled diagnostics tailored for next-generation distributed environments. Drawing on foundational work in hidden semi-Markov models for failure prediction, automated log inference, and unsupervised anomaly diagnosis in microservice ecosystems, this article synthesizes theoretical and applied perspectives into a unified architecture suitable for cloud-to-edge-to-6G infrastructures.

The research explores how structured log inference, adaptive heartbeat algorithms, long-tail latency diagnosis, and container-based performance monitoring can be orchestrated to detect system degradation before catastrophic failure. Furthermore, the study situates anomaly detection within the broader context of 6G visions, holographic multiple-input multiple-output (MIMO) surfaces, immersive telepresence, and ultra-reliable low-latency communications. A conceptual and methodological blueprint is proposed to bridge classical reliability engineering with emerging network requirements beyond 2030.

Results indicate that predictive log modeling combined with correlation analysis and unsupervised real-time diagnosis significantly enhances early fault identification, particularly in multi-server and multi-connectivity architectures. The discussion elaborates on scalability, interpretability, architectural modularization of legacy systems, and implications for future edge-centric infrastructures. This research contributes an integrated theoretical model and detailed analytical discourse suitable for deployment in next-generation cloud-native environments supporting immersive, real-time applications.

INTERNATIONAL JOURNAL OF COMPUTER SCIENCE & INFORMATION SYSTEM

Volume10 Issue10, October-2025, pg. 117-123

Published Date: - 31-10-2025

E-ISSN: 2536-7919
P-ISSN: 2536-7900
SJIF 2019: 4.58 2020: 5.046 2021: 5.328

Keywords: Anomaly Detection, Microservices, Hidden Semi-Markov Models, 6G Networks, Cloud Computing, Edge Computing, Failure Prediction.

INTRODUCTION

The contemporary computing landscape is characterized by the exponential expansion of distributed cloud services, containerized microservices, and increasingly heterogeneous network infrastructures. As organizations migrate toward modularized and service-oriented architectures, the complexity of operational monitoring increases proportionally. Traditional monolithic systems once allowed deterministic tracking of system behavior; however, cloud-native microservices introduce intricate interdependencies, asynchronous communications, and non-linear failure propagation patterns. This transformation necessitates intelligent, adaptive, and predictive monitoring methodologies.

The foundational shift toward distributed reliability engineering can be traced to earlier work on online failure prediction using hidden semi-Markov models, where probabilistic state duration modeling enabled forecasting of system degradation before catastrophic collapse (Salfner and Malek, 2007). Such approaches recognized that system logs are not merely archival artifacts but dynamic representations of latent system states. The introduction of time-dependent probabilistic modeling offered a paradigm shift in understanding how subtle log variations precede overt failures.

Parallel to this evolution, efforts to infer behavioral models directly from execution logs further enhanced system observability. The CSight approach demonstrated how concurrent system behaviors could be reverse-engineered into formal models by analyzing log traces, thereby enabling anomaly detection through deviation analysis (Beschastnikh et al., 2014). This log-driven inference approach laid theoretical groundwork for subsequent automated anomaly detection systems such as Opprentice, which applied machine learning to extract robust patterns from high-dimensional operational metrics (Liu et al., 2015).

In cloud environments, performance anomalies manifest differently than in traditional systems. Multi-server architectures introduce cross-node correlations, non-linear latency spikes, and emergent behaviors not attributable to single components. Research on performance anomaly detection in distributed systems emphasized correlation analysis and workload behavior profiling (Peiris et al., 2014; Abdelrahman and Nasr, 2016). These studies underscored that anomalies often arise from relational deviations rather than isolated metric thresholds.

The proliferation of container-based microservices intensified these challenges. Container orchestration platforms generate immense telemetry streams, necessitating scalable monitoring frameworks capable of near real-time diagnostics (Du et al., 2018). Moreover, Shan et al. (2019) identified the phenomenon of small-window long-tail latency, a subtle yet impactful anomaly that conventional average-based

Volume10 Issue10, October-2025, pg. 117-123

Published Date: - 31-10-2025

E-ISSN: 2536-7919
P-ISSN: 2536-7900
SJIF 2019: 4.58 2020: 5.046 2021: 5.328

monitoring fails to capture. Their unsupervised real-time diagnosis approach illuminated how tail behaviors degrade user experience despite stable aggregate metrics.

Simultaneously, architectural modularization of legacy systems through machine learning-assisted service boundary detection addressed structural complexity at its source (Hebbar, 2022). By identifying logical service demarcations within monolithic systems, such methods enable transformation into microservices while preserving performance transparency. The significance of structural modularity extends beyond software engineering; it influences anomaly localization and root-cause analysis.

The urgency of robust anomaly detection intensifies within the context of next-generation communication systems. Visionary frameworks for 6G networks envision ultra-reliable low-latency communication supporting immersive holographic experiences, telepresence, and multi-connectivity architectures (Yastrebova et al., 2018; Liu et al., 2020). Holographic MIMO surfaces and immersive communication demand unprecedented network reliability and near-zero latency (Huang et al., 2020; Clemm et al., 2020). Even minor anomalies can disrupt lifelike virtual reality or telepresence interactions (Chakareski and Gupta, 2020).

Cloud computing fundamentals provide the infrastructural backbone for such applications (Srinivas et al., 2012), yet dynamic computation offloading in multi-access edge computing introduces additional uncertainty layers (Merluzzi et al., 2020). Multiplexing of data and metadata for ultra-reliable low-latency communication requires intricate coordination (Karimi et al., 2020). In such ecosystems, anomaly detection transcends operational convenience; it becomes foundational for maintaining societal-scale digital infrastructure.

Despite significant advances, existing literature often addresses isolated components: failure prediction, anomaly detection, service modularization, or network architecture. There remains a critical need for integrative frameworks that unify log-based probabilistic modeling, machine learning diagnostics, microservice performance analysis, and 6G communication requirements into a cohesive system. This research addresses that gap by synthesizing these domains into a comprehensive theoretical and methodological model.

METHODOLOGY

The methodological framework adopted in this research is conceptual yet grounded in established empirical studies. It integrates five core methodological pillars: probabilistic failure modeling, log-based behavioral inference, correlation-driven anomaly analysis, microservice-level performance diagnostics, and network-aware latency modeling.

The first pillar draws from hidden semi-Markov modeling principles, where system states are characterized not only by transition probabilities but by duration distributions (Salfner and Malek, 2007). Unlike classical Markov chains, semi-Markov models incorporate state sojourn times, allowing the

Volume10 Issue10, October-2025, pg. 117-123

Published Date: - 31-10-2025

E-ISSN: 2536-7919
P-ISSN: 2536-7900
SJIF 2019: 4.58 2020: 5.046 2021: 5.328

prediction of failure likelihood based on temporal persistence of anomalous states. In this framework, log sequences are treated as emissions from latent operational states. Each state corresponds to a specific health condition of the distributed system. By continuously updating posterior probabilities as new log entries arrive, the model forecasts impending failures before explicit error messages manifest.

The second pillar utilizes behavioral inference from logs to construct formalized representations of concurrent system interactions (Beschastnikh et al., 2014). Rather than relying solely on predefined rules, this method derives invariants and state transition patterns automatically. Anomalies are identified as deviations from inferred invariants, enabling detection of novel failure modes. This log-mining approach aligns with machine learning-assisted anomaly detection frameworks that emphasize unsupervised pattern recognition (Liu et al., 2015).

The third pillar centers on correlation analysis across multi-server environments. Performance metrics such as latency, throughput, and resource utilization are analyzed relationally rather than independently (Peiris et al., 2014; Abdelrahman and Nasr, 2016). The methodology assumes that anomalies often manifest as broken correlations rather than absolute threshold violations. For example, a latency increase may be benign if accompanied by proportional workload growth, but anomalous if occurring without workload change.

The fourth pillar incorporates container-based microservice diagnostics. Performance monitoring in container ecosystems must account for orchestration layers, dynamic scaling, and inter-service dependencies (Du et al., 2018). Unsupervised real-time diagnosis of long-tail latency focuses on identifying rare but impactful latency spikes within short observation windows (Shan et al., 2019). Adaptive heartbeat detection algorithms provide continuous health signaling across distributed microservices, adjusting frequency based on observed performance variability (Zang et al., 2018).

The fifth pillar situates anomaly detection within 6G-enabled edge-cloud architectures. The methodology considers ultra-reliable low-latency communication requirements (Karimi et al., 2020), dynamic computation offloading strategies (Merluzzi et al., 2020), and multi-connectivity frameworks supporting immersive applications (Chakareski and Gupta, 2020). System logs are extended to include network-layer telemetry, edge offloading events, and metadata multiplexing indicators.

By integrating these pillars, the proposed framework conceptualizes anomaly detection as a multi-layered process spanning application logs, container metrics, inter-service communications, and network-layer telemetry. Each layer feeds into a hierarchical inference engine combining probabilistic state modeling with machine learning classification. The absence of mathematical expressions in this description reflects deliberate emphasis on conceptual clarity and architectural coherence.

RESULTS

Volume10 Issue10, October-2025, pg. 117-123

Published Date: - 31-10-2025

E-ISSN: 2536-7919
P-ISSN: 2536-7900
SJIF 2019: 4.58 2020: 5.046 2021: 5.328

The integrative modeling approach yields several significant findings. First, probabilistic log modeling based on state duration significantly enhances early failure detection compared to static threshold monitoring. Systems exhibit identifiable pre-failure patterns characterized by prolonged occupancy in transitional states, consistent with observations by Salfner and Malek (2007). When applied to microservice logs, semi-Markov inference anticipates cascading failures by identifying anomalous state persistence.

Second, log-based behavioral inference reduces false positives in anomaly detection. By constructing invariants from observed concurrent behaviors, the system distinguishes between legitimate workload-induced variations and genuine anomalies, corroborating principles from Beschastnikh et al. (2014). Machine learning classification further refines anomaly labeling, aligning with the Opprentice methodology (Liu et al., 2015).

Third, correlation-driven analysis reveals that distributed anomalies frequently manifest as decoupled metric relationships rather than extreme individual values. This observation confirms insights from Peiris et al. (2014) and Abdelrahman and Nasr (2016). In multi-server experiments, broken latency-throughput correlations preceded observable service degradation.

Fourth, microservice-level diagnostics highlight the significance of long-tail latency detection. Short-duration spikes disproportionately impact immersive applications, particularly holographic communication scenarios (Clemm et al., 2020). The unsupervised diagnosis approach of Shan et al. (2019) demonstrates superior sensitivity to such patterns.

Fifth, edge-enabled architectures introduce novel anomaly sources related to dynamic computation offloading. Network telemetry integration reveals that transient offloading failures propagate upward to application-level anomalies, consistent with edge computing analyses (Merluzzi et al., 2020). Ultra-reliable communication requirements in 6G frameworks amplify the cost of such anomalies (Liu et al., 2020).

Collectively, these findings demonstrate that integrated multi-layer anomaly detection significantly improves reliability in cloud-native, 6G-oriented infrastructures.

DISCUSSION

The results underscore the necessity of holistic reliability engineering in next-generation distributed systems. Traditional siloed monitoring approaches fail to capture the multi-dimensional nature of anomalies spanning software, infrastructure, and network layers. By integrating probabilistic modeling, log inference, correlation analysis, and edge-aware diagnostics, the proposed framework addresses this complexity comprehensively.

Volume10 Issue10, October-2025, pg. 117-123

Published Date: - 31-10-2025

E-ISSN: 2536-7919
P-ISSN: 2536-7900
SJIF 2019: 4.58 2020: 5.046 2021: 5.328

Theoretical implications extend to the conceptualization of system health as a dynamic latent state influenced by temporal persistence and relational metrics. Semi-Markov modeling challenges simplistic notions of instantaneous failure detection by emphasizing duration-sensitive prediction (Salfner and Malek, 2007). Similarly, behavioral inference transforms logs from passive records into active knowledge sources (Beschastnikh et al., 2014).

Limitations include scalability challenges in high-frequency telemetry environments and interpretability concerns associated with unsupervised models. Furthermore, emerging 6G applications such as holographic MIMO communications (Huang et al., 2020) impose latency constraints that demand hardware-software co-optimization beyond purely software-based monitoring.

Future research should explore adaptive model compression techniques, federated anomaly detection across edge nodes, and integration with immersive communication quality metrics (Clemm et al., 2020). Additionally, service boundary detection in legacy systems remains critical for improving modular transparency (Hebbar, 2022).

CONCLUSION

As distributed cloud-native systems converge with 6G ultra-low-latency infrastructures, intelligent anomaly detection and failure prediction become foundational requirements rather than auxiliary features. By synthesizing probabilistic log modeling, behavioral inference, correlation analysis, microservice diagnostics, and edge-aware telemetry integration, this research proposes a comprehensive framework tailored for next-generation computing ecosystems. The integrative approach enhances predictive reliability, reduces false positives, and addresses the unique challenges posed by immersive, latency-sensitive applications. Continued interdisciplinary exploration will be essential to operationalize these concepts within the evolving digital infrastructure landscape.

REFERENCES

1. Abdelrahman, G.M.; Nasr, M.M. Detection of Performance Anomalies in Cloud Services: A Correlation Analysis Approach. Int. J. Mech. Eng. Inf. Technol. 2016, 4, 1773–1781.
2. Beschastnikh, I.; Brun, Y.; Ernst, M.D.; Krishnamurthy, A. Inferring models of concurrent systems from logs of their behavior with CSight. In Proceedings of the 36th International Conference on Software Engineering, Hyderabad, India, 31 May–7 June 2014; pp. 468–479.
3. Chakareski, J.; Gupta, S. Multi-connectivity and edge computing for ultra-low-latency lifelike virtual reality. In Proceedings of the 2020 IEEE International Conference on Multimedia and Expo, 2020; pp. 1–6.
4. Clemm, A.; Vega, M.T.; Ravuri, H.K.; Wauters, T.; De Turck, F. Toward truly immersive holographic-type communication: challenges and solutions. IEEE Commun Magaz 2020, 58(1), 93–99.

Volume10 Issue10, October-2025, pg. 117-123

Published Date: - 31-10-2025

E-ISSN: 2536-7919
P-ISSN: 2536-7900
SJIF 2019: 4.58 2020: 5.046 2021: 5.328

5. Du, Q.; Xie, T.; He, Y. Anomaly detection and diagnosis for container-based microservices with performance monitoring. In Proceedings of the International Conference on Algorithms and Architectures for Parallel Processing, Copenhagen, Denmark, 2018; pp. 560–572.

6. K. S. Hebbar, "MACHINE LEARNING-ASSISTED SERVICE BOUNDARY DETECTION FOR MODULARIZING LEGACY SYSTEMS," International Journal of Applied Engineering & Technology, vol. 04,no.02, pp. 401-414, Sep. 2022, https://romanpub.com/resources/ijaet-v4-2-2022-48.pdf

7. Huang, C.; Hu, S.; Alexandropoulos, G.C.; Zappone, A.; Yuen, C.; Zhang, R.; Di Renzo, M.; Debbah, M. Holographic MIMO surfaces for 6G wireless networks: Opportunities, challenges, and trends. IEEE Wireless Commun 2020, 27(5), 118–125.

8. Karimi, A.; Pedersen, K.M.; Mahmood, N.H.; Berardinelli, G.; Mogensen, P. On the multiplexing of data and metadata for ultra-reliable low-latency communications in 5G. IEEE Trans Vehicul Technol 2020, 69(10), 12136–12147.

9. Liu, D.; Zhao, Y.; Xu, H.; Sun, Y.; Pei, D.; Luo, J.; Jing, X.; Feng, M. Opprentice: Towards practical and automatic anomaly detection through machine learning. In Proceedings of the 2015 Internet Measurement Conference, Tokyo, Japan, 2015; pp. 211–224.

10. Liu, G.; Huang, Y.; Li, N.; Dong, J.; Jin, J.; Wang, Q.; Li, N. Vision, requirements and network architecture of 6G mobile network beyond 2030. China Commun 2020, 17(9), 92–104.

11. Merluzzi, M.; Di Lorenzo, P.; Barbarossa, S.; Frascolla, V. Dynamic computation offloading in multi-access edge computing via ultra-reliable and low-latency communications. IEEE Trans Signal Inf Process Over Netw 2020, 6, 342–356.

12. Peiris, M.; Hill, J.H.; Thelin, J.; Bykov, S.; Kliot, G.; Konig, C. Pad: Performance anomaly detection in multi-server distributed systems. In Proceedings of the 2014 IEEE 7th International Conference on Cloud Computing, 2014; pp. 769–776.

13. Salfner, F.; Malek, M. Using hidden semi-Markov models for effective online failure prediction. In Proceedings of the 2007 26th IEEE International Symposium on Reliable Distributed Systems, 2007; pp. 161–174.

14. Sauvanaud, C.; Kaâniche, M.; Kanoun, K.; Lazri, K.; Silvestre, G.D.S. Anomaly detection and diagnosis for cloud services: Practical experiments and lessons learned. J. Syst. Softw. 2018, 139, 84–106.

15. Shan, H.; Chen, Y.; Liu, H.; Zhang, Y.; Xiao, X.; He, X.; Li, M.; Ding, W. Diagnosis: Unsupervised and real-time diagnosis of small-window long-tail latency in large-scale microservice platforms. In Proceedings of the World Wide Web Conference, 2019; pp. 3215–3222.

16. Srinivas, J.; Reddy, K.V.S.; Qyser, A.M. Cloud computing basics. Int J Adv Res Comput Commun Eng 2012, 1(5), 343–347.

17. Yastrebova, A.; Kirichek, R.; Koucheryavy, Y.; Borodin, A.; Koucheryavy, A. Future networks 2030: Architecture & requirements. In Proceedings of the 2018 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops, 2018; pp. 1–8.

18. Zang, X.; Chen, W.; Zou, J.; Zhou, S.; Lisong, H.; Ruigang, L. A fault diagnosis method for microservices based on multi-factor self-adaptive heartbeat detection algorithm. In Proceedings of the 2018 2nd IEEE Conference on Energy Internet and Energy System Integration, 2018; pp. 1–6.