UDC 004.415.53:004.8

STRATEGIC TRANSFORMATION OF QUALITY ASSURANCE: FROM MANUAL TESTING TO AUTONOMOUS AI-DRIVEN ANALYSIS AND INTELLIGENT ROOT CAUSE ANALYSIS

Bogatov V.O., Poliakov A.O.
e-mail: vitalii.bohatov@nure.ua, andrii.poliakov@nure.ua
Kharkiv National University of Radio Electronics,
Department of Applied Mathematics
Kharkiv, Ukraine

У роботі здійснено системний аналіз процесу трансформації контролю якості програмного забезпечення в умовах зростаючої складності розподілених архітектур. Досліджено перехід від мануальних методів перевірки до використання великих мовних моделей (LLM) для автоматизації пошуку першопричин дефектів (Root Cause Analysis) на основі системних логів. Окрему увагу приділено методологіям підвищення точності діагностики за допомогою агентних підходів та інтеграції ймовірнісного виведення через баєсівські мережі.

The modern software industry is undergoing a paradigm shift in Quality Assurance (QA). As cloud computing and microservices generate terabytes of unstructured data, manual log analysis has become physically impossible. From a systems analysis perspective, distributed systems are complex objects where a single defect causes nonlinear, cascading failures, often masking the true root cause. Traditional manual testing has exhausted its scalability, necessitating intelligent decision support systems to transform raw event streams into structured knowledge.

Large Language Models (LLMs) are central to this transformation. Their semantic understanding of technical language allows for a fundamentally new approach to log processing. Research by Akhtar, Khan, and Parkinson identifies six dominant AI categories in this field: anomaly detection, error monitoring, semantic parsing, root cause analysis (RCA), security management, and threat identification. Unlike classical algorithms like Drain or Spell, LLMs generalize effectively, detecting anomalies in previously unseen scenarios and adapting to dynamic log structures without needing rigid regular expressions. The transition to AI-driven QA requires new training methodologies, such as fine-tuning, Retrieval-Augmented Generation (RAG), and in-context learning. Analysis [1] shows that while baseline ReAct architectures offer only 35.5% diagnostic accuracy, the Flow-of-Action method – combining standard operating procedures with autonomous AI agents – raises accuracy to 64.01%. This proves that effectiveness depends on integrating the model with the operational context and architectural interdependencies. Agent-based systems allow AI to move beyond passive text analysis to actively querying monitoring tools and verifying hypotheses.

Automating RCA in microservices is now a critical intersection of testing and architecture. Mhatre and Sheikh [2] demonstrate solutions using Kafka clusters for intelligent error classification, distinguishing business logic errors from system failures with over 90% accuracy. Fine-tuned models analyzing stack traces can generate diagnostic reports in under one second, significantly reducing Mean Time to Recovery (MTTR). Here, the systems analyst acts as a data flow architect, filtering "noise" to ensure the neural network receives relevant incident data.

Further development involves hybrid frameworks combining LLMs with classical statistical methods. Pedroso [3] proposes deploying dynamic Bayesian networks alongside LLMs in Kubernetes/Istio environments. This provides a probabilistic assessment of component influence, adding interpretability to "black-box" AI decisions. Testing with Chaos Mesh and Locust confirms these hybrid systems remain reliable even under high-load conditions [3].

Practical industrial applications, such as the ABB case described by Komonen [4], highlight the transition from manual verification to automated failure summarization. RAG architectures allow models to utilize "knowledge chains" from documentation and incident history, helping testers distinguish genuine product defects from "flaky" infrastructure-related tests. This systemic approach enables a move from reactive correction to predictive quality modeling.

In summary, integrating AI into QA is the only viable path for complex IT systems. LLMs overcome complexity barriers previously requiring high-level experts, reducing MTTR by up to 80% [5]. Systems analysis is evolving toward autonomous agent ecosystems that handle evidence collection and defect classification, freeing engineers for strategic design. Future research will focus on data privacy and mitigating hallucinations through formal validation and RAG technologies.

References:

1. Bogatov V. Use of large language models for event log analysis and root cause analysis in distributed systems // Студентська наукова конференція здобувачів вищої освіти та молодих вчених «Сталий розвиток сучасних інформаційних технологій»: матеріали конференції. Харків : ХНАДУ, 2025. С. 133–136.

2. Mhatre A., Shaikh A. LLM for Automated Root Cause Analysis in Microservices Architectures. Journal of Information Systems Engineering and Management. 2025. Vol. 10, No 13. P. 449-460. URL: https://doi.org/10.52783/jisem.v10i13s.2100

3. Pedroso, D. F., Almeida, L., Pulcinelli, L. E. G., Aisawa, W. A. A., Dutra, I., & Bruschi, S. M. Anomaly detection and root cause analysis in cloud-native environments using large language models and bayesian networks. TechRxiv. 2025. P. 1-15. URL: https://doi.org/10.36227/techrxiv.174016565.57888427/v1

4. Komonen A. AI-Powered Log Analysis for Test Automation. Aalto University. Espoo, 2025. 75 p. URL: https://aaltodoc.aalto.fi/bitstreams/76111c77-b2f3-4cfb-8acb-a9ec699a0265/download

5. Cut MTTR by 50%: How AI-Powered Root Cause Analysis is Revolutionizing Incident Response. DEV Community. 2025. URL: https://dev.to/devactivity/cut-mttr-by-50-how-ai-powered-root-cause-analysis-is-revolutionizing-incident-response-2n7b (дата звернення: 09.03.2026).