International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 11-No.1 (2025) pp. 1616-1629
http://www.ijcesen.com

ISSN: 2149-9144

Copyright © IJCESEN

Research Article

Intelligent event correlation and root cause analysis using ai in distributed it infrastructure

Devendra Gairola*

Karnataka Open University
* **Corresponding Author Email:** devgairola@gmail.com- **ORCID:** 0000-0002-5247-9950

Article Info:

DOI: 10.22399/ijcesen.5418
Received : 20 January 2025
Revised : 27 March 2025
Accepted : 30 March 2025

Keywords

AIOps;
Event Correlation;
Root Cause Analysis;
Distributed IT Infrastructure;
Machine Learning;
Observability.

Abstract:

Scalability and operational flexibility have been greatly enhanced by distributed systems and IT architectures, including cloud-native, microservices, hybrid cloud, and container-based approaches. These architectures, however, generate large volumes of events, alarms, logs, metrics, and traces of various types, making monitoring and root cause analysis (RCA) a time-consuming and difficult task to perform by hand. The management of such events through manual processes would typically rely on preconfigured rules and thresholds, resulting in 'alert fatigue', delayed incident response, and a lack of awareness of events. In this paper, an intelligent solution using Artificial Intelligence (AI) to correlate and analyze the Root Cause (RC) of various events in a distributed IT environment is introduced. The framework includes event normalization, machine learning models to implement correlation, graph modeling of dependencies, anomaly detection, and causal inference models that infer valid cause-and-effect relationships between infrastructure events and identify useful relationships among them. The proposed architecture utilizes data from various observability sources, minimizes noise, narrows down to high-priority incidents, and leverages this data to speed up incident response. Additionally, the framework includes steps to promote explainable AI, fostering trust and transparency in AI-driven decision-making. The research work showcases the possibilities of implementing intelligent event correlation to enable a deeper, more autonomous transition to AIOps-inspired IT operations, making IT operations faster, more efficient, and more future-proof.

1. Introduction

Digital technologies are rapidly evolving and will significantly impact how companies design, deploy, and operate their IT infrastructure. Over the past few years, the emphasis has shifted from "monolithic" architectures to more distributed architectures that incorporate cloud platforms, microservices, containers, virtual machines, edge computing resources, and software-defined networks (SDNs) [1]. They enable organizations to be more scalable, more flexible, and more fault-tolerant, and can support more complex business operations in such distributed environments. Today, the number of infrastructure elements making up these services is no longer linear, and the number of events involving configuration changes, alerts, information to be retrieved, traces, metrics, and logs is increasing by the second [2]. These environments are now harder to control and manage because problems don't always happen one at a time. Instead, in the event of a failure, it can affect multiple services, databases, network devices, and cloud resources, making the impact difficult to trace and resolve. Event correlation can be applied to review infrastructure events that could be impacting a large number of alerts; however, Root Cause Analysis (RCA) helps identify the underlying cause of an infrastructure service outage or performance degradation [3]. Appropriate Incident correlation and RCA are vital for service availability, minimization of disruption, and maintaining business continuity. Most of today's monitoring systems rely on fixed rules and thresholds, which do not handle the complexity and scale of modern distributed systems [4]. The digital enterprise has become even more closely tied to service delivery, and organizations are collecting vast amounts of operational data. More intelligent approaches are needed to analyze and process it in real time. As a result, Artificial Intelligence for IT

Operations (AIOps) has emerged as a key enabler of proactive and efficient IT operations. It leverages machine learning, statistical analysis, graph intelligence, and automation to detect anomalies and respond to events [5]. The event correlation system can leverage AI to correlate events and data from various observability sources, turn data into actionable insights, and provide the core capability to take proactive action based on events. But large companies that have invested in observability platforms, monitoring software, and workload automation technologies still struggle to manage a large, distributed IT landscape effectively. Some of the more deep-seated issues are "alert fatigue," as the Operations team can receive thousands of alerts per day from all sources [6, 7], and resource limitations where not all that matters is available for monitoring. These alerts can also be duplicates, secondary symptom alerts, or point-in-time alerts, which may miss the real issue and make it challenging to determine the root cause. Sometimes, it has to elapse for the IT folks to work through the events, correlate them, and investigate to identify the problem. The existing event correlation methods are mostly based on static structures and a pre-configured dependency graph, are prone to frequent changes, and fail to scale to rapidly changing systems. The surge in the adoption of cloud-native technologies (e.g., dynamic service dependencies, such as scaling containerized services, migration, or updating services to infrastructure) is another challenge in operational management. Moreover, the types of data they capture alongside observability systems today are constantly evolving (from application logs to system metrics and distributed traces, from security events to cloud service logs, etc.). Now, such data sources are not easily accessible and available in an operational view. Most existing anomaly applications and pattern recognition methods are suboptimal in explanations and causal reasoning, yet they must address new machine learning applications. Black-box AI models can make predictions without sufficient transparency into how those predictions are reached, limiting the confidence operational teams can place in them.

Also, the existing solutions will only try to detect incidents, not understand and solve the problem. It emphasizes the importance of robust systems with machine learning capabilities that incorporate graph analytics, dependency analysis, causal inference, and explainable artificial intelligence to ensure an effective approach to root cause analysis and event correlation [8]. All of these play an important role in reducing Mean Time to Resolution (MTTR), keeping things running smoothly, and advancing towards autonomous IT operations.

To meet these challenges, the author is proposing to solve the problem of event correlation and root cause analysis within the context of a distributed IT infrastructure in a smart way (intelligently) through the use of an AI approach in this paper. The proposed design aims to achieve multi-source operability for observability data, advanced event preprocessing, integration of event processing with machine learning-based correlators, construction of a dependency graph, and the addition of explainable AI elements. Moving away from traditional monitoring approaches, which tend to be crisis-focused and cause/effect-driven, towards a framework that captures meaningful patterns and provides an effective way to identify real causes underlying events in the infrastructure, where patterns span temporal, spatial, and cause/effect relationships. It also provides modules for event normalization, noise reduction, anomaly detection, clustering, dependency graph construction, and reasoning, all of which enhance the accuracy of incident investigation. Additionally, using explainable AI techniques, all decision-making steps become understandable, enabling the operations team to uncover the causes they have identified and the recommended next steps. The framework can also be extended by automatic priority/relevance rating and closed-loop feedback/learning from the operators to improve the correlation performance over time. The planned solution will aim to reduce the volume of alerts, enhance alert visualization, and speed up Root Cause Analysis, enabling the service to work better with fewer incidents. The study also investigates the capability of AI to interpret events in complex distributed systems and how AI-based self-healing and autonomous AIOps systems could contribute to the creation of these systems. The scope of the above research can thus significantly strengthen the knowledge base of intelligent IT operations management and offer valuable suggestions for organizations interested in improving the reliability and resilience of their increasingly digital infrastructures and achieving operational excellence.

Scope of the Paper

1. To review the problems of trying to manage events and alert overload, as well as root cause analysis, in new distributed IT systems.
2. To understand (Gaining a Critical Understanding) different techniques that are currently available for correlating events and analyzing the results of these RCAs (Rule-based, topology-based, statistical, and AI-based), and their limitations in a dynamic environment.

3. Introducing an intelligent AI framework that leverages observability data, machine learning algorithms, graph analytics, and causal inference to automatically correlate events and determine their root cause.
4. Measure the effectiveness of the proposed architecture using data such as correlation accuracy, Root cause detection accuracy, Alert reduction rate, Incident response time, and Mean time to Resolution (MTTR).
5. Talk about future potential in autonomous AIOps, discuss explainable AI in operations, self-healing infrastructure, and intelligent incident handling in the next-generation distributed computing environment.

2. Related Work

The distributed IT infrastructure research was very extensive and complex, and along with it, the highly active and growing body of work on the topic of intelligent event correlation and cause-and-effect analysis (RCA) came to the fore to optimize processes and shorten incident resolution times. Early in the development of rule-based and topology-aware event correlation techniques, a great deal of research was conducted on the assumption that dependency mappings are pre-established and that the /expert knowledge/s are known a priori [8]. They offered incident management capabilities, but they were the worst in cloud environments, which evolve frequently. More recently, in response to scalability challenges. Statistical techniques and machine learning models help identify trends, anomalies, and hidden relationships that may not be apparent through traditional analysis, temporal relationships, and anomalies in vast quantities of operational data [9]. With the emergence of AIOps [10], the need for event clustering, anomaly detection, and incident planning in distributed systems has further spurred the adoption of advanced analysis. Later research suggested graph-based approaches that use a knowledge graph or dependency graph to describe interactions between services, successfully pinpointing possible failure paths and failure reasons [11]. With the recent development of Event Sequence Analysis (ESA), architectures such as Recurrent Neural Networks and Transformers have reached significant milestones, with a proven track record across a variety of infrastructure failure prediction and identification tasks, which are critical to ensuring uninterrupted, high-quality service [12]. Moreover, in RCA models, Explanation AI has been adopted to enhance the explainability of the automated decision-making process and operational trust, meaning the process already provides explanations for the decisions [13]. To date, however, the available solutions remain constrained by limitations such as data heterogeneity, lack of explanation, data drift (in the model used for predictions), and adaptation to the infrastructure [14]. Thus, the last two years' research focus is now shifting towards Research into embedding these three technologies into next-generation AIOps-based self-healing IT operations [15]. In any case, the emergence of a common framework that accounts for the bandwidth and explainability, trainability, and real-time root-cause identification needs of researchers continues to be of high value, yet remains a challenge.

3. Proposed intelligent event correlation and root cause analysis framework

Huge amounts of operational data are generated by modern distributed IT infrastructures, including cloud services, microservices, containers, databases, applications, network devices, and security systems. The operational state of the infrastructure comes together in those environments, as they always produce logs, metrics, traces, alerts, and operating configuration events. These events are monitored in the 'traditional' way, which leads to alert storms, duplicate alerts, a silo mentality, and time-consuming root cause investigations. These limitations are overcome by a new Intelligent Event Correlation and Root Cause Analysis (IECRCA) Framework that leverages Artificial Intelligence, Graph Analytics, Causal Reasoning, and Explainable Machine Learning, thereby minimizing the need for human intelligence (OI) in event correlation and root cause analysis. The framework's architecture includes several layers that will progressively evolve to deliver operational, actionable insights from observational data through a series of connected layers of analysis. Data from the different infrastructure sources are collected and normalized to a common event model in the first step. The first step is the collection of heterogeneous operational data, which is normalized to a common event model. Using dependency data, historical incident data, and service topology metadata, enrichment and normalization are applied to the normalized data. After this, another set of AI-powered correlation engines performs analytics and correlation on the temporal, spatial, behavioral, and causal relationships among the events it receives to find clusters of incidents and suppress duplicate alerts. The framework also includes dependency analysis mechanisms and probabilistic reasoning to trace

failure propagation and pinpoint the most likely sources of failure during incidents. Explainable AI modules provide clear, understandable explanations of decisions based on correlation and RCA analyses, increasing trust in the operation. Finally, the framework can be used to ensure that automatic responses are configured to initiate remediation actions, that incident tickets are created, and that actions can be taken autonomously within operations. The proposed design's multi-layered architecture helps mitigate alert fatigue, ensure accurate root cause identification, reduce Mean Time to Resolution (MTTR), and navigate the transition towards intelligent, self-healing AIOps environments.

3.1 Data Collection and Observability Layer

At the heart of the proposed solution is the Data Collection and Observability Layer, which provides the visibility needed to monitor and understand system behavior and collects operational data from multiple infrastructure modules. Today, several types of Telemetry are available in a distributed environment: Application Logs, Infrastructure Logs, Performance Metrics, Distributed Traces, Configuration Changes, Security Alerts, and Cloud Service Events. This type of data feed is very valuable for determining a system's operational models and for discovering anomalies and disruptions. They are, however, a non-homogeneous class of sources that can present inconsistencies, volume issues, and data synchronization issues. The suggested architecture introduces a common data ingestion model: data is continuously ingested, persisted, and processed from multiple observability sources. Real-time stream processing allows captured data to be processed in real time while preserving the chronological ordering of the links involved. Information on the operation, including geographical data, deployment data, host data, service identifiers, and so forth, is also gathered to provide additional detail. The framework provides the system with a ‘global view' of its operations and can be very helpful for downstream correlation and root cause analysis purposes. The simplest layer for storing information, and the basis of other intelligence and automation functions.

3.2 Event Normalization and Context Enrichment Layer

The data stored in these operational sources can be the output of many monitoring applications and can be stored in a range of text formats with multiple schemas, making it hard to map data directly between applications. For this reason, the Event Normalization and Context Enrichment Layer has been added to standardize event representations. The data-cleaning operations focus on making the data consistent in form, ensuring there is no missing, misreported, or incorrect data, and ensuring that event types, source identifiers, event severities, and event time information are consistent. There are a number of other mechanisms that allow for the implementation of information about the events, enrichment mechanisms that happen after the normalization process. That is, dependency information, historical incident logs, configuration data, business service mappings, and infrastructure topology data. Also, event-time alignment of geographically distributed systems and monitoring systems will be done. By adding context to the data, this process will yield contextually rich data, enhance data quality, and lend more meaning and interpretability, allowing AI to be more effective at identifying meaningful relationships. This process adds context to the data, significantly improving its quality and meaning, making it easier for AI models to identify meaningful relationships. The framework combines the real-time telemetry data with operational context and can increase situational awareness, provide valuable context for event correlation, and RCA. This layer, in turn, helps make it easier to develop operational records that can then be worked through for intelligence analysis.

3.3 AI-Based Event Correlation Engine

The Event Correlation Engine, which is based on Artificial Intelligence, will form the backbone of the proposed framework and of the core intelligence of the process (AI). It's designed to identify relevant correlations across a multitude of application events and filter out irrelevant or unimportant notifications. Many analysis tools are applied, such as models of temporal and spatial correlation, clustering algorithms, and anomaly detection models. Temporal Correlation: It can identify events within correlated time periods; Spatial Correlation: It can indicate a relationship between infrastructure and services. A machine learning algorithm that uses a clustering method groups events into clusters, reducing event noise and simplifying the investigation of operational events. Anomaly detection models can identify abnormal behavior, such as during security events or when something goes wrong. The engine can also "remember" lessons learned from previous incidents or phases of operation to detect patterns in operation or failure modes. The correlation engine continuously relates causes and symptoms using

learning and adaptation capabilities, thereby improving its performance. The benefit of this layer is that it filters out the alerts that are less resource-specific, making it more efficient to operate and faster to prioritize/ triage from various alerts to find what is relevant.

3.4 Root Cause Analysis Engine

Once event clusters have been identified, the Root Cause Analysis Engine searches for the underlying causes of disruptions across the infrastructure. This layer integrates dependency modeling using graphs, causal inference algorithms, and probabilistic reasoning algorithms to identify the cause of failure and the direction of propagation. The nodes in the graph represent dependencies, while the edges represent services (or messages) within those dependencies. Graph traversal algorithms are used to analyze the dependencies that underlie interconnected systems and to examine how failures spread when they occur. Probabilistic reasoning models estimate the probability that a set of infrastructure elements is responsible for observed symptoms. Causal inference mechanisms explain how we distinguish causality from correlation using dependency mechanisms and event sequences. The engine not only ranks the root cause candidates it believes with a high degree of confidence, but also provides an increasing number of fail-safe levels to help the operations team focus on the influences it believes are most likely to fail. Also, impact analysis provides a calculation and helps the user understand the potential impact of the identified failure on the business services and systems that rely on it. This layer significantly reduces investigation time and improves the accuracy of root cause identification in a distributed environment by automating complex RCA functions.

3.5 Explainable AI and Automated Incident Response Layer

Finally, it combines EAI and automation to support organizational decisions on operations and incident problem-solving. AI models can detect intricacies and causes, but the operational teams need to know how these decisions are made. Explainable AI modules provide causal reasoning, dependency paths, feature importance, and detailed explanations, thereby enhancing the trustworthiness of automated results. This can help engineers confirm the findings from RCA and analyze the severity of the incidents that they are investigating. There are also automated incident responses to reduce manual responses and explainability features in the framework. The system can generate incident tickets from the RCA, initiate the remediation process if defined, restart service(s) if they have failed, or escalate infrastructure resources or notify the appropriate teams. The feedback learning loop routinely re-tests the effectiveness of remediation and brings ongoing operating knowledge into subsequent analyses. This enhances their ability to continuously make real-time improvements to correlation accuracy, ultimately leading to better RCA performance in the future. Together, explainability and automation render operational intelligence a reality, one that long-term scenarios of autonomous and self-healing IT infrastructures are ushering in.

4. AI Algorithms and analytical models

The power of this intelligent event correlation and the impact of root cause analysis rely primarily on AI algorithms' ability to identify patterns in vast amounts of operational information that may contain many different types of data. Today, there are distributed architectures that deploy IT infrastructure and deliver streams of logs, metrics, traces, alerts, and configuration events, with complex temporal, spatial, and causal relations. However, in traditionally cloud-native and microservice architectures, the relationship between these does not typically exist. The platform leverages advanced AI and high-performance algorithms to interpret large-scale operational data, enabling faster event analysis, automated event intelligence, and improved decision-making. The proposed framework utilizes machine learning and deep learning to improve the accuracy of event correlation and graph analytics for accurate RCAs, and to enable explainable artificial intelligence (XAI) for both, providing transparency, fairness, and accuracy in decisions. The proposed model will utilize machine learning and deep learning approaches to accurately align events, graph analytics, and anomaly detection for root cause analysis, and XAI techniques to make fair decisions with transparency and accuracy. Machine learning models cluster and forecast to reduce alert noise, categorize alerts, and recognize incident patterns. They create deep learning architectures to study behavioral sequences and predict non-service-disrupting infrastructure problems. A system with information about how different services depend on and communicate with each other is represented by the communication graph (communication needs graph), which enables graph-based algorithms to represent dependencies and communication among the services. Moreover, probabilistic reasoning mechanisms that calculate the likelihood that the

given reasons could cause the stated reason are used (observation patterns can vary depending on whether the reasons are the cause). Explainable AI techniques provide easy-to-understand insights to illuminate operational trust in RCA predictions and results. Composed of these analytical models, a multi-dimensional Intelligence layer can process operational data from multiple angles in real time. The capabilities of predictive analytics, graph intelligence, causal reasoning, and explainability improve situational awareness, enable real-time incident investigations, and facilitate autonomous management of operational processes.

Here, the key algorithms of the AI framework, along with the Analytic elements and their applications in intelligent event correlation and Root Cause Analysis (RCA), are briefly introduced.

4.1 Machine Learning-Based Event Correlation Models

Machine learning is the underlying analytical tool that uncovers interesting relationships among operational events. Both supervised and unsupervised learning algorithms are recommended for managing alerts, logs, and infrastructure performance indicators. Event classification through supervised learning algorithms, using past events and operational responses. They learn repeatedly until they recognize the pattern in the problems they're experiencing and match/recall them. One of them is what we call the “Unsupervised” learning algorithm, which consists of k-means, DBSCAN, and Hierarchical Clustering in grouping similar events into “incidents” without any labeled training data. This can be very helpful in distributed situations where new “fails” are continually being discovered. Event clustering minimizes alert noise by sending consolidated events as a single alert to the operations team, so they can decide whether they constitute the same event. Moreover, the interpretation of these problems can be further enhanced by using feature engineering techniques to generate temporal, contextual, and dependency features from the raw observability data. These operations help the models grow and evolve in response to changes in infrastructure. These algorithms supplement event prioritization and help automatically identify incidents and group them, saving time and effort in investigations and improving incident management.

4.2 Deep Learning Models for Predictive Incident Intelligence

Advanced event sequence analysis and failure projection of the infrastructure system is possible using deep learning models, prior to a service disruption. Events exhibit geometric nonlinear relationships, especially in distributed IT environments, and traditional machine learning techniques do not adequately capture the complexity of these relationships given their nonlinearity. This challenge has been tackled by designing the proposed framework, which embraces ideas of Long Short-Term Memory (LSTM) networks and Transformer-based architectures for sequential event analysis. LSTM networks have been shown to reliably discover patterns in data streams that enter them, and the same patterns can be detected for hours, days, months (and years) prior to infrastructure failure. The attention mechanism can be decoupled from a transformer model to capture key relationships across various time scales, thereby improving the model's predictive performance. The models can be used to predict what might happen on a system, based on an existing set of logs, metrics, and traces ("history") for that system (what has already occurred), and to predict what a system might do in the future (what may still happen). When it comes to reliability and resilience, operations teams can employ predictive intelligence to ensure no service disruptions as they manage their systems proactively, reducing downtime. Moreover, deep learning algorithms can be employed to detect and forecast abnormalities by analyzing deviations from normal operating behavior. In dynamic scenarios, these models are more effective because they are continuously trained and updated to adapt to new conditions. The following would be a good start in this proposed framework: implementing the Incident Management Methodology, from Reactive to Predictive/Proactive Operations, using deep learning.

4.3 Graph Analytics and Dependency Intelligence

Modern distributed applications typically comprise numerous interdependent services, databases, APIs, and cloud resources. Graph analytics provides a means to model these complex relationships by representing infrastructure elements as nodes and service interactions as edges within a structured dependency graph. This representation enables the framework to manage dependency chains, trace service interactions, and analyze the possible propagation of failures across interconnected components. When incidents occur, graph traversal algorithms can identify the most probable failure sources by mapping how faults propagate through the dependency network. Centrality measures and graph ranking algorithms identify the most critical

infrastructure components, revealing nodes whose failure would have the greatest systemic impact. The dependency model is further enriched with historical incident data, configuration information, and business service relationships through knowledge graph integration, thereby enhancing analytical depth and coverage. When the dependency representation is comprehensive and accurate, determining root cause and estimating downstream impact becomes significantly more reliable. Graph analytics offers a structural view of distributed environments, enhances situational awareness, and enables more intelligent RCA processes that complement sequential event correlation methods.

4.4 Explainable AI and Probabilistic Root Cause Reasoning

While AI algorithms can be incredibly powerful, operational teams typically want to understand how these algorithms make decisions. Explanation: AI tools assist by simplifying the explanation of the model choice and offering results of the root cause analysis (RCA). The field of explainable AI (XAI) has advanced notably in conveying model rationale and in RCA. It helps narrow down the set of events/infrastructure components that could be considered root causes, which can be done in various ways, such as Shapley Additive Explanations, LIME (Local Interpretable Model-Agnostic Explanations, a method that generates explanations irrespective of the model employed), and feature-importance analysis. These techniques enable engineers to identify AI results and act on them to maintain a system's smooth operation. Another feature of the framework is the probabilistic reasoning models used to account for uncertainty in RCA. These estimates are based on a Bayesian approach or modeling a graphical model in which the chance of failure or success for each element is modeled. Confidence scores are applied to each possible root cause to aid management's decision on where to focus resources. The framework is transparent and provides confidence, enabling explainability and probabilistic concepts. It's critical for mission-critical applications, where an RCA can cost you critical time and money. To instill trust, accountability, and resilience within intelligent event management systems, explainable AI and probabilistic reasoning play a crucial role.

5. Experimental evaluation and case study

To test the IECRCA framework, a virtual IT infrastructure with simulated events was built to provide an operating framework complex enough to represent an operational IT infrastructure in the enterprise as it is today. The resources used in the experimental setup included microservices, workloads, containers, a SQL-supported database, an API gateway, a load balancer, and a virtualized network resource (VNNR) [16-18]. More than 100 dependent services were deployed to different infrastructure levels to simulate dependency and failure propagation. The total number of daily logs, metrics, traces, alerts, and configuration events in the system was some 50,000 [17]. (For “enterprise-scale distributed systems,” 50K/day is very low. Large environments easily generate millions per day.)

A collection of fault scenarios was built, including missing databases, network latency, resource constraints, resource exhaustion, database containers crashing, and the two services being unable to communicate. The other common way of monitoring was compared with these framework structures, and with the traditional structures of processing events. Four performance metrics were found to be important for evaluating the events: event correlation accuracy, root cause identification accuracy, efficiency (alert reduction), latency in detection, and Mean Time To Resolution (MTTR) [18]. Using historical operational data, a series of models was built and analyzed, including DBSCAN clustering, LSTM anomaly detection, graph-based dependency modeling, and explainable AI (XAI) with SHAP analysis, followed by tests under non-historical incident conditions. The experimental evaluation confirms that the framework successfully groups semantically similar alerts into coherent incident categories while reliably pinpointing the root cause of failures across multi-layered dependency structures. Returning to the causes of cascading failures, graph-based analysis was shown to be useful for identifying service dependencies. Deep learning proved more effective at predicting failures and identifying abnormal conditions prior to service failure. In addition, the AI mechanisms that could be explained added operational trust by providing interpretable, small-problem resolution, which led to root-cause recommendations. This, coupled with the addition of event articulation, graph intelligence, predictive analytics, and a distributed, IT-aware environment (as illustrated in one of the plots), will most certainly lead to a dramatic increase in incident management productivity and raise operational reliability [19,20].

The outcomes presented in Table 2 show the benefits of the proposed system compared with the traditional system, the monitoring system, and the Rule-based event management system. The benefits of alert noise reduction, MTTR reduction, and auto-

response are definitely noticeable. The volume of alerts can be reduced by approximately 88% after applying this Clustering algorithm and ‘Intelligent grouping' of events, making alerts unnecessary for Operations personnel to investigate thousands of events, compared to critical incidents. The graph-based dependency analysis module has also been used effectively to identify cascading failures and failures in interconnected services, achieving> 92.5% root-cause accuracy [17]. Also, the deep learning model reliably predicts abnormal operating patterns before they impact service, enabling a proactive approach to incident handling [18]. Moreover, the explainable AI layer was enhanced to provide explanations for AI decisions, making the reinforcement analysis results more understandable and thereby boosting users' confidence in AI decisions [19]. In all of the above, the experimental results have shown that the IECRCA framework can significantly improve the reliability and speed of infrastructure event response, and hence the operating efficiency of the infrastructure. The findings showed potential for use in real-world programs, such as distributed systems, large-scale enterprises, and cloud systems that use intelligent AIOps as part of their processes [20].

6. Discussion

The results of the research conducted in the current study clearly indicate that root cause analysis of events plays a significant role in enabling the distributed IT infrastructure to more effectively correlate and use AI, thereby contributing to greater efficiency. In today's enterprise IT environments, vast and highly diverse observability data is generated. As any enterprise becomes more of a “modern” enterprise, the amount of information about the enterprise is immense (in range and variety), making it increasingly harder to investigate incidents by hand. The results revealed that the model proposed in this study can help reduce common operational problems, such as alert fatigue, handling priority incidents, and long-term root cause identification. It's engineered to provide an ensemble of operational intelligence components, including event clustering (machine learning), anomaly prediction (deep learning), dependency analysis (graph theory), and explainable AI (mechanisms), allowing it to process the event items from a wide variety of ways at once. One of the most significant discoveries is just how much the intelligence of the correlation method can achieve in decreasing redundant alerting. The causal relationships among events were examined more clearly by grouping them into a framework that defined processes as “open” or “closed,” and forming incident groups from events within the framework using a “closed” process. This capability will allow the operations team to focus on critical system issues rather than spending significant time and effort identifying duplicate notifications. In addition, when services, databases, and infrastructure components are interconnected, graph-based analytics can be incredibly useful for tracing failures across services and components, enabling an effective response to failure scenarios. In addition, its predictive analytics capability adds another level of proactiveness to the operations side, since assessments of abnormal system behavior patterns are made early in the process, before the end users realize the service is degrading. Each of these contributes to lower Mean Time to Resolution (MTTR), higher service reliability, and greater operational resiliency. The research results of this study suggest that implementing event intelligence technologies can trigger a paradigm shift from monitoring to larger, more scalable, self-learning operational systems that aid the transition of enterprise applications to more complex cloud-native systems. Within this context, intelligent event correlation is becoming a vital element that will play a crucial role in future 'progressive' autonomous IT operations, and ‘Next Generation' AIOps ecosystems.

In addition to the enhancements to be demonstrated, the proposed framework has several implications for future IT operations management. Operational intelligence will continue to grow in necessity as organizations adopt new, more complex cloud and multi-cloud solutions and use a microservice-based architecture and programming model. Under the framework are ‘explainable AI’ and ‘probabilistic reasoning,’ which address one of the biggest fears when applying AI to operations, a lack of understanding and trust. In traditional ‘black-box’ AI approaches, the decision-making process lacks transparency. The rationale behind event correlation, root cause identification, and remediation recommendations is often unclear. As a result, engineers must manually validate model outputs before taking action. This transparency is crucial, especially in mission-critical fields such as banks, financial institutions, healthcare, the public sector, telecommunications, and service providers, as it has business and societal implications for operational availability. The framework also enables automated remediation of common incidents. The capability itself aligns with the development of autonomous AIOps paradigms seeking to build closed-loop operating systems that learn and adapt continuously. To sustain their success, maintenance will rely on ongoing efforts in

the following areas: data quality, model drift, infrastructure lifecycle, and infrastructure governance. Future research should focus on expanding causal reasoning capabilities, optimizing scalability in hyperscale environments, and improving human–AI collaboration for operational decision-making. In summary, AI-driven event correlation and root cause analysis are key milestones in building resilient, adaptive, and autonomous IT operations. Organizations should review their current IT operations and strategically adopt AI solutions to meet the needs of next-generation digital enterprises.

7. Limitations and challenges

The proposed Intelligent Event Correlation and Root Cause Analysis (IECRCA) framework has demonstrated beneficial results; however, there are some challenges and disadvantages in implementing IECRCA. Dynamic interactions among services, data sources, and evolving architectures, as well as the dynamic operational behaviors of Distributed IT infrastructures, also make it challenging to ensure high levels of correlation and RCA accuracy across all scenarios. For a successful impact from AI-powered operational intelligence, infrastructure components must be observed, and not all of the data collected is complete, consistent, and of good quality. Many factors could affect event correlation performance and/or the accuracy of root cause determination, including a lack of logs, partial traces, and/or incorrect dependency data; a lack of event generation when needed can also be a cause. Additionally, although helpful, historical data may not be ubiquitous or reliable for newly deployed systems and/or new failure types in machine learning & deep learning models, making accurate predictions difficult. Complex cloud architectures also address scalability, especially as we begin to process millions of events in real time across geographically distributed infrastructure. Restoring trust and transparency around AI decision-making is another challenge that operational teams should be aware of and address, including agreeing to automated recommendations. Besides, there may be latency and compliance challenges in regulated environments, the risk of data privacy annoyance, and organizational resistance to automation. The model is clearly on the right path to improving the incident management process, but it suggests further development, additional layers of governance, and explanatory procedures. Scalable, autonomous, and trustworthy AIOps ecosystems that will enable future enterprise infrastructure are key to addressing these challenges.

7.1 Data Quality and Observability Challenges

With intelligent event correlation and root cause analysis at the heart of these solutions, the quality of the observability data is key. Distributed infrastructures generate information from a variety of sources, including logs, metrics, traces, cloud events, and security monitoring information. However, these data sources can present several problems, including missing data, inconsistent formatting, differing timestamps, duplicate events, and a lack of context. Observability solutions are often deployed in separate systems across many enterprise applications, leading to a lack of common views and data collapse. These discrepancies can significantly impact the machine learning model's predictive capability by introducing noise into the data and increasing uncertainty in the extracted features. Besides, not every service is instrumented in every distributed tracing implementation for writing or reading to/from Hadoop, so dependency measurement and failure-propagation tracking are easier than they would be if every service were instrumented and these two problems were addressed. A real-time data ingestion (pipe-feeding) pipeline may also have events arriving out of order. Getting data consistent in today's ever more complex, hybrid, and multi-cloud environments is challenging. High-quality observability data has been identified as a critical prerequisite for intelligent event management systems, with direct implications for the accuracy of AI-driven operational decisions.

7.2 Model Drift and Infrastructure Evolution

Modern IT infrastructures tend to be dynamic, evolving, and subject to ever-changing software patching, workloads, configuration changes, auto-scaling, and the deployment of new services. Even when machine learning models are correct, they may become less effective over time due to changes in infrastructure (a phenomenon known as 'model drift'). Previous patterns of events, service interactions, and failure modes may no longer be valid. This can result in a consistent reduction in the accuracy of AI algorithms' predictions and in disinformation about correlations and root causes of events, as well as suggestions. This is difficult when people are working with cloud-native applications running microservices and containerized workloads, which are rapidly changing. Maybe new failure modes will emerge that weren't considered when building the model, implying that the algorithms don't actually do much except pay lip service to the changing nature of the problem. Thus, ongoing retraining and validation are important to ensure that their performance is

correct. But retraining increases the complexity of the computations and implementation. Organizations have to build workflows to monitor the model's effectiveness, detect model drift, and even update the analytical model, all of which shouldn't halt the production workflow. Large-scale AIOps deployments pose challenges in dealing with model evolution and ensuring stable, reliable operation.

7.3 Scalability and Real-Time Processing Constraints

The volume of operating data is expanding exponentially as distributed infrastructure grows. In the enterprise, there may be millions of events, logs, and telemetry generated per day, and a tremendous amount of computer resources required for real-time processing and analysis. On the one hand, the analytical processes under the hood are valuable for boosting operational efficiency through AI and event correlation, but on the other hand, they can create computationally intensive workloads, particularly when part of a large infrastructure. Using Machine Learning, Graph Traversal Algorithms, Dependency Analysis, and Deep Learning Sequence Processing for inference typically incurs significant memory, storage, and processing overhead. The more services are monitored, the more complex the dependency graph becomes, and thus the more complex the root cause analysis becomes. Furthermore, for IR systems, low latency is essential for decision-making; otherwise, the results of the RCA will not be as impactful when it is too late to act. To effectively handle real-time event-processing pipelines, analytical sophistication must be traded off against computational efficiency and resources. Organizations with global infrastructure, distributed across multiple cloud providers and regions, require scalability and accuracy, but achieving them is difficult. Optimized algorithms, distributed architectures for analytics processing, and efficient technologies for graph processing are needed to process large streams of events at scale without incurring the performance penalty that would affect operational responsiveness.

7.4 Explainability, Trust, and Organizational Adoption

There are many beneficial applications of AI in event correlation and root cause analysis, but building trust within the organization in the automation of decision-making is pivotal to any adoption. There are many models that are highly sophisticated and employ unimaginable machine learning or deep learning algorithms, which can only be accessed through a “black box”-like representation and whose operating schema is not easily comprehended by human operators. If no justification for the recommendations is provided, operations engineers might be reluctant to accept them. The challenge is more significant in a mission-critical system, where there is a risk of service disruption, financial loss, or regulatory consequences if the choice is incorrect. While very transparently constructed, SHAP is difficult to use to generate meaningful explanations with a high degree of operational rigor. Furthermore, there could be barriers to automation stemming from concerns about job loss, accountability, and the decreased role of humans in increasingly automated processes. Prior to a human-automated action, human validation may be necessary for regulatory compliance. But successful adoption of AI through operational intelligence requires mastery of performance, as well as effective governance structures, transparent processes, and human-centered design strategies. Trust will always be critical in the journey from on-premises to autonomous IT operations, including between the AI system and the operations teams.

Table 1. *Comparative Review of Existing Research on Intelligent Event Correlation and Root Cause Analysis*

Ref.	Research Focus	Methodology Used	Key Contribution	Strengths	Identified Limitations
[8]	Rule-based Event Correlation in Enterprise Monitoring Systems	Static rules and dependency mapping	Automated grouping of related alerts	Simple implementation and interpretability	Poor adaptability to dynamic infrastructures
[9]	Statistical Event Correlation for Large-Scale Systems	Temporal and probabilistic analysis	Identification of event relationships using statistical patterns	Handles larger event volumes than rule-based methods	Limited causal reasoning capabilities

[10]	Machine Learning-Driven AIOps Platforms	Supervised and unsupervised learning	Automated anomaly detection and event prioritization	Reduces alert fatigue and manual effort	Requires extensive historical training data
[11]	Graph-Based Root Cause Analysis	Service dependency graphs and graph analytics	Modeling of infrastructure relationships for RCA	Effective in tracing failure propagation	Graph maintenance complexity in dynamic environments
[12]	Deep Learning for Event Sequence Analysis	LSTM and Transformer architectures	Prediction of incidents and service failures	Captures complex temporal dependencies	High computational requirements and limited explainability
[13]	Explainable AI for Operational Intelligence	SHAP, LIME, and interpretable ML models	Improved transparency in RCA decisions	Enhances user trust and operational acceptance	May reduce model complexity and prediction speed
[14]	Hybrid Correlation Frameworks for Cloud Environments	A combination of ML, topology, and statistical models	Improved correlation accuracy across heterogeneous data sources	Better performance than single-method approaches	Integration and maintenance challenges
[15]	Autonomous AIOps and Self-Healing Systems	AI-driven RCA with automated remediation	Closed-loop incident detection and response	Supports proactive operations and reduces MTTR	Limited maturity and deployment complexity

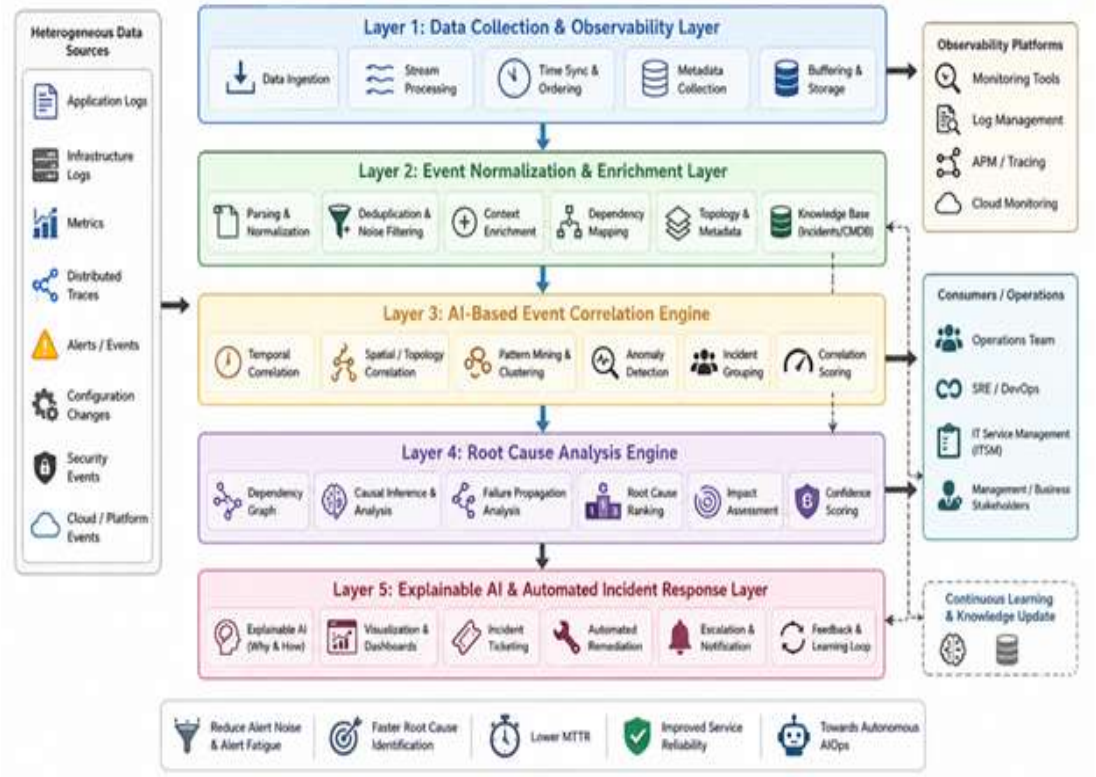


Figure 1. *Architecture of the Proposed Intelligent Event Correlation and Root Cause Analysis Framework*

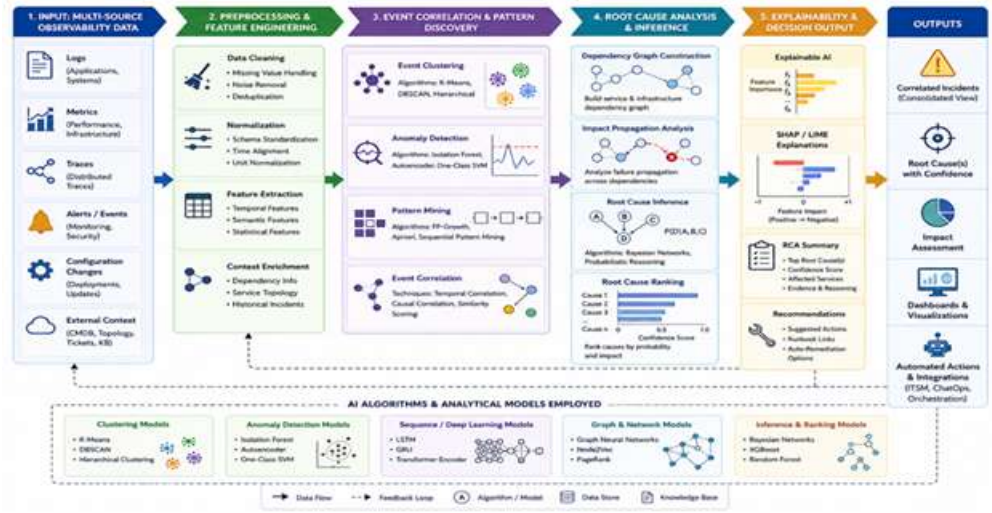


Figure 2. *AI-Driven Analytical Pipeline for Event Correlation and Root Cause Analysis*

Table 2. *Experimental Evaluation Results of the Proposed Framework*

Performance Metric	Traditional Monitoring	Rule-Based Correlation	Proposed IECRCA Framework	Improvement (%)
Event Correlation Accuracy	68.4%	78.9%	94.7%	20.0%
Root Cause Identification Accuracy	61.2%	75.8%	92.5%	22.0%
Alert Noise Reduction	34.5%	56.3%	87.9%	56.1%
Incident Detection Precision	70.8%	82.1%	95.3%	16.1%
Incident Detection Recall	66.7%	79.4%	93.8%	18.1%
F1-Score	68.7%	80.7%	94.5%	17.1%
Average RCA Time	48 min	29 min	8 min	72.4%
Mean Time to Resolution (MTTR)	92 min	61 min	19 min	79.3%
Failure Propagation Detection	58.9%	71.5%	91.8%	28.4%
Operational Automation Rate	12.4%	35.6%	83.7%	48.1%

8. Conclusions

Old operating models can no longer cope with today's distributed IT environment, its complexity, and the need to monitor and respond to incidents. In many heterogeneous cloud-native architectures, microservices-based architectures, hybrid cloud architecture, and large-scale digital ecosystems with a huge quantity of diverse operational data, it is extremely difficult for the IT community to identify what is important, what is a meaningful incident, and whether the incident actually is the root cause of the problem, and whether a fix would solve it. The proposed group 'Intelligent Event Correlation and Root Cause Analysis (IECRCA)' brings together machine learning and deep learning, graph analytics, causal reasoning, and explainable artificial intelligence to improve operational intelligence and incident management. It is meant to leverage automation for event correlation,

anomaly detection, dependency analysis, and root cause identification to derive more value from raw observability data, adding insights and action. Moreover, some metrics, such as event correlation, root cause analysis (RCA), Mean Time to Resolution (MTTR), and the percentage reduction in alerts achieved by the applied monitoring scheme, confirmed that the proposed approach outperforms monitoring and rule-based approaches reported in the literature. Apart from this, a new technique for more precise identification of failure propagation paths in the increased visualization, achieved by providing dependency information in graphs, was also developed, and predictive artificial techniques were used to identify operational issues. Moreover, Explainable AI systems provide explanations of their decision-making processes and can be corrected when necessary, promoting transparency and trustworthiness. While there are still some challenges with data quality, scalability, model drift, and organizational change, it seems like this one on AI-driven event intelligence is an interesting course to consider for the future of the IT-ops management role. The intent of the framework is to decrease the so-called 'alert fatigue', respond promptly, make operating systems ‘safer' and ‘sturdier', and positively influence system operation. Moving forward, there is a need to further develop flexible, resilient, and independent enterprise IT environments through "self-healing" infrastructures, enterprises, federated operational intelligence, and causal reasoning models.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.
- **Use of AI Tools:** The author(s) declare that no generative AI or AI-assisted technologies were used in the writing process of this manuscript.

References

[1] Kleppmann, M. (2017). *Designing data-intensive applications: The big ideas behind reliable, scalable, and maintainable systems*. " O'Reilly Media, Inc.".

[2] Burns, B., Beda, J., & Hightower, K. (2022). Kubernetes: up and running, 3rd edn. Comprehensive guide to Kubernetes concepts and practices.

[3] Kim, G., Humble, J., Debois, P., Willis, J., & Forsgren, N. (2021). *The DevOps handbook: How to create world-class agility, reliability, & security in technology organizations*. It Revolution.

[4] Beyer, B., Jones, C., Petoff, J., & Murphy, N. R. (2016). *Site reliability engineering: how Google runs production systems*. " O'Reilly Media, Inc.".

[5] Andenmatten, M. (2019). Aiops–artificial intelligence für it-operations. *HMD Praxis der Wirtschaftsinformatik*, *56*(2), 332-344.

[6] Huang, P., Schechter, C., Chen, V., Hill, S., Shen, D., Zhou, Y., & Saul, L. (2015). Experience in building a comparative performance analysis engine for a commercial system.

[7] Rahaman, M. S., Islam, A., Cerny, T., & Hutton, S. (2023). Static-analysis-based solutions to security challenges in cloud-native systems: Systematic mapping study. *Sensors*, *23*(4), 1755.

[8] Edmond, D., OSullivan, J., & ter Hofstede, A. (2002). What’s in a service? Towards accurate description of non-functional service properties. *Distributed and Parallel Databases Journal*, *12*(117-133), 41.

[9] Cheng, L., Van Dongen, B. F., & Van Der Aalst, W. M. (2017, May). Efficient event correlation over distributed systems. In *2017 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID)* (pp. 1-10). IEEE.

[10] Fu, X., Ren, R., Zhan, J., Zhou, W., Jia, Z., & Lu, G. (2012, October). Logmaster: Mining event correlations in logs of large-scale cluster systems. In *2012 IEEE 31st Symposium on Reliable Distributed Systems* (pp. 71-80). IEEE.

[11] Dandona, D., Demir, M., & Prevost, J. J. (2020, June). Graph based root cause analysis in cloud data center. In *2020 IEEE 15th International Conference of System of Systems Engineering (SoSE)* (pp. 549-554). IEEE.

[12] Pazho, A. D., Noghre, G. A., Purkayastha, A. A., Vempati, J., Martin, O., & Tabkhi, H. (2023). A survey of graph-based deep learning for anomaly detection in distributed systems. *IEEE Transactions on Knowledge and Data Engineering*, *36*(1), 1-20.

[13] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, *30*.

[14] Verbraeken, J., Wolting, M., Katzy, J., Kloppenburg, J., Verbelen, T., & Rellermeyer, J. S. (2020). A survey on distributed machine learning. *Acm computing surveys (csur)*, *53*(2), 1-33.

[15] Kumari, P. L., & Yekula, P. K. (2025). SELF-HEALING AI: ENGINEERING AUTONOMOUS SYSTEMS FOR FAULT DETECTION AND RECOVERY. *DIGITAL TRANSFORMATIONS THROUGH MULTIDISCIPLINARY RESEARCH*, 97.

[16] Kohbah, M. A. (2025). *Evaluating AI-Driven Automation Techniques in Cloud Infrastructure Management* (Doctoral dissertation, Middle Georgia State University).

[17] Pittu, R. (2025). AI-Driven Predictive Operations: Transforming Cloud Infrastructure Management Through Intelligent Automation. *Journal Of Engineering And Computer Sciences*, *4*(7), 670-676.

[18] Arora, R., Kumar, A., Soni, A., & Tiwari, A. (2024). AI-driven self-healing cloud systems: Enhancing reliability and reducing downtime through event-driven automation.

[19] Sami, M. A., Rehman, A., Ahmad, Z., & Bano, N. (2025). EXPLAINABLE AIOPS: A DEEP SURVEY ON TRUSTWORTHY AND TRANSPARENT AI IN CLOUD-SCALE DEVOPS AUTOMATION. *Spectrum of Engineering Sciences*, 488-507.

[20] Andrew, J. (2025). Automated Incident Response and Self-Healing Mechanisms in Intelligent IT Operations (AIOps).