Alarm Root Cause Analysis using AI/ML in MSO Networks

A technical paper prepared for presentation at SCTE TechExpo24

Jonathan Kwan, PhD, PEng
Product Line Manager
Fujitsu Network Communications, Inc.
jonathan.kwan@fujitsu.com

Table of Contents

List of Figures

1. Introduction

Next-generation networks are increasingly challenging to manage as service touchpoints are continuously added. Coupled with the need for legacy networks to work alongside new network rollouts, operators are seeing a growing number of "alarm storms" generated by all these systems and services. These alarm storms not only extend the time needed to evaluate issues, but also make it more challenging to balance the resources used to investigate issues and manage networks.

Automation, artificial intelligence (AI) and machine learning (ML) in network operations is increasingly popular with multiple system operators (MSOs) as a means to reduce costs, predict network performance, and drive network efficiencies. By using test data representative of an MSO network to train neural networks alongside a classification engine, the relationship between nodes and grouping of like behavior was explored. This paper will show how AI/ML techniques were successfully implemented to suppress 99% of the alarms, locate and partition the root cause of an alarm storm with high accuracy (first recommendation accuracy up to 80%), and reduced time-to-solution (from hours to minutes), resulting in higher customer satisfaction and network reliability.

2. Challenges Faced by MSOs

As aforementioned, management of next-generation networks is becoming more challenging due to the increased number of service touchpoints and the need for legacy networks to work alongside new technologies. Consequently, when network issues arise, thousands of alarms or more can be generated within a short period of time; an event known as an "alarm storm". Complex, interconnected, multigenerational, multivendor networks make troubleshooting very challenging for network operations center (NOC) staff. In fact, tracing the root cause of an alarm storm can take hundreds of staff hours.

Furthermore, depending on the number of devices in an operator's network, incidents that occur per year, average time to address the issue, and compensation related to SLAs, alarm storms can incur millions of dollars in direct costs [1], [2], [3], not to mention indirect costs related to customer satisfaction and reputation. Customers now have more options than ever before for connectivity and services, and will switch to another provider if they receive unreliable service, causing expensive churn for MSOs.

At a time where customers are demanding faster, more reliable service, while many MSOs are simultaneously facing flat average revenue per unit (ARPU), how can an operator improve network reliability while reducing costs? And how can MSOs balance resources to manage their networks and investigate issues? One way to achieve this is for the operations team to reduce mean time to repair.

Currently, element management systems (EMS) or controllers use rules defined by the operator or equipment vendor to categorize events. However, the network and the relationship between its components is actively evolving. What worked yesterday may not work well tomorrow. Subject matter experts are needed to update rules and event policies, which can be time-consuming and inconsistent regarding quality. Furthermore, alarm filtering and correlation remove useful data from the analysis, which can lead to misdiagnosis.

Automation and AI/ML in network operations can be an invaluable tool to reduce costs, predict network performance, and drive operational efficiencies.

3. Solution Approach

In this paper, we propose designing an AI/ML app that performs root cause analysis (RCA) and placing it in an MSO's network to work alongside NOC staff, as shown in Fig. 1.

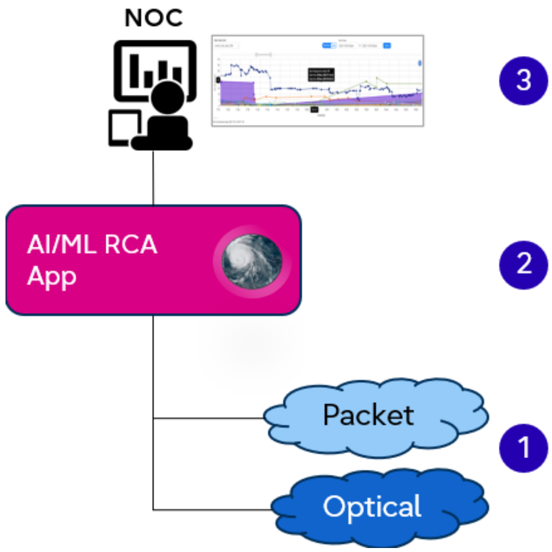

Figure 1 - AI/ML for RCA in a NOC

When an incident occurs in the network, an alarm storm occurs with up to thousands of alarms, labeled (1) in Fig. 1. Initially, the AI/ML RCA app must answer the following two questions:

1) *What* is the problem that is causing this alarm storm?
2) *Where* is the problem that is causing this alarm storm?

The proposed AI/ML RCA app labeled (2) in Fig. 1 answers these two questions by recognizing the event and beginning analysis.

One common analytical approach is correlation. There has been a lot of activity around correlation in recent times. One research challenge faced by academia and industry alike is that correlation is good to build an understanding of patterns and relationships incrementally, but it is not necessarily good at identifying causations or resolutions. From an MSO network's perspective, there may be alarm profiles based on various similarities among events. Conclusions are drawn on how interrelated these alarm profiles and events are over time. However, it is challenging to consistently pinpoint the correct root cause using correlations alone, since correlation does not necessarily mean causation.

Beyond simple correlation, this paper proposes the use of neural networks in conjunction with a classification engine in the AI/ML RCA app to execute highly accurate and fast root cause analysis, as shown in Fig. 2. The classification engine automatically discovers the relationship between nodes and build groups of like behavior from the fault data. Understanding behavior groups assists in training the neural network to distinguish between noise and meaningful alarm groups. Furthermore, to best suit the needs of MSOs, the proposed AI/ML RCA app is network-centric, multivendor, multilayer, and scalable, as well as easy to implement and use. A problem that can be solved technically will also need to have low barriers to adoption for widespread uptake by MSOs.

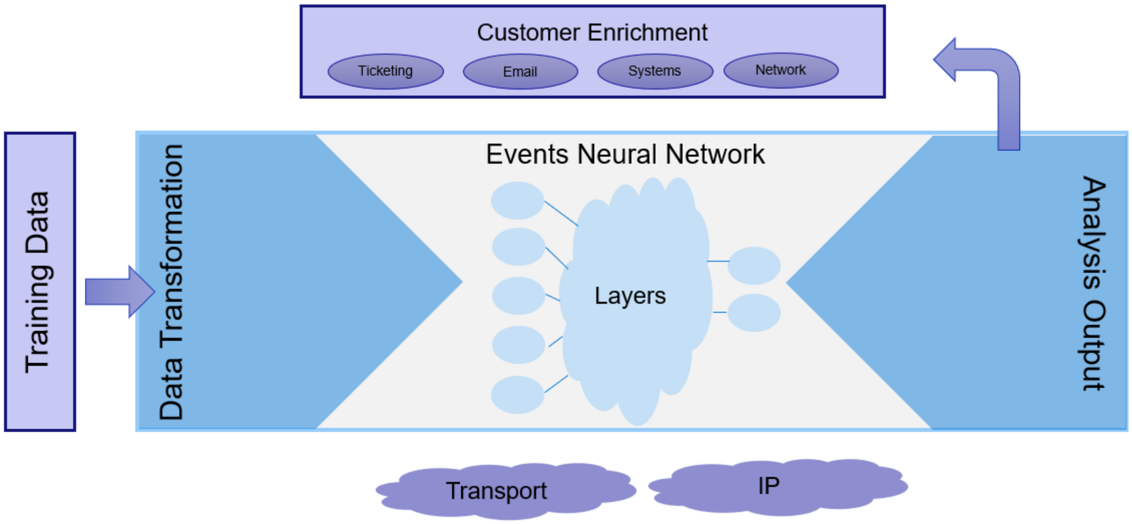

Figure 2 - AI/ML RCA application neural network interface

After delivering accurate and fast root cause analysis, the AI/ML RCA app identifies not only the root cause, but also informs NOC staff in plain language with network events knowledge baked in. For example, the app tells the staff, "I am 90% confident the problem is a fiber issue. Here is the port identifier (ID) and the node ID."

To further enhance accuracy, user solution tagging is supported. User solution tagging allows NOC staff to easily confirm the AI/ML RCA app's recommendations within the app's user interface, further enhancing the app's AI/ML model to further enhance its accuracy and time-to-solution over time.

Combining all the above, an AI/ML RCA app is designed using the architecture shown in Fig. 3.

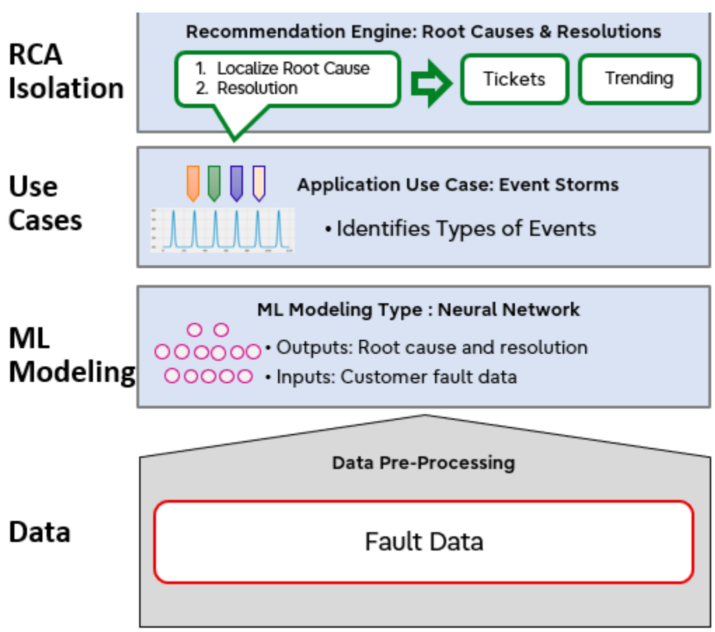

Figure 3 - AI/ML RCA application high-level design

4. Test Results

To test the results and effectiveness of this solution approach, an AI/ML RCA app was built and tests were conducted on test data representative of an MSO network.

Figure 4 - Alarm storms example

Fig. 4 depicts a series of alarm storm windows. The AI/ML RCA app's first challenge is to distinguish whether all the alarms are one storm or multiple storms, and which ones are noise. The AI/ML RCA app's classification engine, alongside the neural network, was able to automatically discover the relationship between nodes and build groups of like behavior from the data.

In this specific example, there were 6,060 raised alarms. The AI/ML RCA app was able to classify them into 47 unique alarms, resulting in a >99% alarm suppression rate. The incident ticket record was 50, which means the AI/ML RCA app achieved a 94% success rate in identifying events. Furthermore, the plain language recommendation system was able to provide the correct root cause in its top 5 recommendations in 41 of the 47 accurately identified events, *i.e.* 87% accuracy. The first recommendation accuracy is 80% based on our test results. Adding user solution tagging provides feedback to the app and will further enhance its accuracy and time-to-solution over time.

In terms of reducing mean time to repair, one example that can be drawn from this test data is that the AI/ML RCA app detected an event on the first day at 17:20 and identified a root cause within minutes. In contrast, the standard event tool did not recognize the outage until the second day at 22:29, since no root cause was identified from the events on the previous day, and the problem was ignored. There was a 29-hour difference between the AI/ML RCA app and the standard event tool.

These results used test data representative of an MSO network and are consistent with implementations on operators' real-world networks.

5. Conclusion

In this paper, AI/ML techniques were shown to have been successfully implemented using an RCA app to suppress 99% of the alarms, locate and partition the root cause of an alarm storm with high accuracy of up to 80%, and to reduce root cause analysis time from hours to minutes. The AI/ML RCA app was able to rapidly cut through environmental interference to locate and partition the root cause of a disruption.

Network behavior was used to automatically adapt neural models and automation workflows, helping manage rapid network behavior changes faster and more accurately. The app's accuracy and time-to-solution can be further enhanced via feedback from NOC staff.

The AI/ML RCA app is a powerful network operations tool that can assist NOC staff of all experience levels to rapidly and accurately make decisions and allow staff to focus on more important tasks rather than engaging with mundane or frustrating processes, saving MSOs time and money. Automation and AI/ML is clearly an invaluable tool that can reduce costs, predict network performance, and drive network efficiencies, resulting in increased customer satisfaction and network reliability for MSOs.

Future research directions include implementation of generative AI for analysis of documentation, such as product manuals, and files, such as log files, to automatically and rapidly generate solution recommendations.

Abbreviations

AI	artificial intelligence
ARPU	average revenue per user
EMS	element management system
ML	machine learning
MSO	multiple systems operator
NOC	network operations center
RCA	root cause analysis
SLA	service level agreement
ID	identifier

Bibliography & References

[1] CTV News, "Rogers to spend $261M to split networks, but can't quantify economic impact of outage." https://www.ctvnews.ca/business/rogers-to-spend-261m-to-split-networks-but-can-t-quantify-economic-impact-of-outage-1.6041829 (accessed Jul. 19, 2024).

[2] The Asahi Shimbun, "KDDI says it will compensate 36 million users for massive outage." https://www.asahi.com/ajw/articles/14682791 (accessed Jul. 19, 2024).

[3] Reuters, "AT&T to credit customers a full day of service for Thursday outage." https://www.reuters.com/business/media-telecom/att-credit-customers-full-day-service-thursday-outage-2024-02-25 (accessed Jul. 19, 2024).