Root Cause Analysis of Poor FTTR Quality Based on Transformer Mechanisms

YU Weichao, LIU Yang, ZHANG Junxiong,
YE Junliang, GE Xiaohu
(School of Electronic Information and Communication, Huazhong University of Science and Technology, Wuhan 430074, China)

DOI: 10.12142/ZTECOM.202504006

https://kns.cnki.net/kcms/detail/34.1294.TN.20251209.1014.002.html, published online December 09, 2025

Manuscript received: 2025-11-11

Abstract: Fiber-to-the-Room (FTTR) has emerged as the core architecture for next-generation home and enterprise networks, offering gigabit-level bandwidth and seamless wireless coverage. However, the complex multi-device topology of FTTR networks presents significant challenges in identifying sources of network performance degradation and conducting accurate root cause analysis. Conventional approaches often fail to deliver efficient and precise operational improvements. To address this issue, this paper proposes a Transformer-based multi-task learning model designed for automated root cause analysis in FTTR environments. The model integrates multidimensional time-series data collected from access points (APs), enabling the simultaneous detection of APs experiencing performance degradation and the classification of underlying root causes, such as weak signal coverage, network congestion, and signal interference. To facilitate model training and evaluation, a multi-label dataset is generated using a discrete-event simulation platform implemented in MATLAB. Experimental results demonstrate that the proposed Transformer-based multi-task learning model achieves a root cause classification accuracy of 96.75%, significantly outperforming baseline models including Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Random Forest, and eXtreme Gradient Boosting (XGBoost). This approach enables the rapid identification of performance degradation causes in FTTR networks, offering actionable insights for network optimization, reduced operational costs, and enhanced user experience.

Keywords: FTTR; root cause analysis; Transformer mechanisms; Wi-Fi; multi-task learning

Citation (Format 1): YU W C, LIU Y, ZHANG J X, et al. Root cause analysis of poor FTTR quality based on Transformer mechanisms [J]. ZTE Communications, 2025, 23(4): 37 – 47. DOI: 10.12142/ZTECOM.202504006
Citation (Format 2): W. C. Yu, Y. Liu, J. X. Zhang, et al., "Root cause analysis of poor FTTR quality based on Transformer mechanisms," *ZTE Communications*, vol. 23, no. 4, pp. 37 – 47, Dec. 2025. doi: 10.12142/ZTECOM.202504006.

1 Introduction

Fiber-to-the-Room (FTTR)[1] is a key framework for next-generation home and enterprise networks, with global deployment accelerating due to its broad application prospects. Built on an all-optical fiber network architecture, this technology delivers ultra-high bandwidth (ranging from 1 Gbit/s to 10 Gbit/s), low latency, high stability, and seamless coverage to end users. It effectively supports bandwidth- and performance-intensive applications such as 4K/8K ultra-high-definition video streaming, cloud gaming, virtual reality/augmented reality (VR/AR), smart home systems, and remote work and education[2]. As global fixed broadband networks progress toward gigabit and ten-gigabit capabilities, the deployment scale of FTTR continues to expand, accompanied by a steady increase in market penetration. This trend underscores FTTR as a critical direction in the evolution of next-generation broadband access technologies.

In December 2020, the European Telecommunications Standards Institute (ETSI) included FTTR in the fifth-generation fixed network (F5G) standard system[3], reflecting international recognition of its technological innovation and standardization progress. Currently, ten-gigabit access solutions that integrate 50G passive optical networks (PON) and FTTR are progressing through pilot testing and commercial deployment stages. Many service providers are actively promoting this technology as a strategic initiative to enhance user experience, capture high-end market segments, and strengthen competitive advantage. Furthermore, as end-users increasingly demand high-quality indoor coverage and stable network performance, achieving high reliability, efficient operation and maintenance, and consistently superior user experiences in FTTR networks has become a central focus for both academic research and industrial development.

The rapid expansion of FTTR networks, coupled with their highly complex topology, which includes primary gateways,

This work is supported in part by the National Key R&D Program of China under Grant No. 2024YFE0200504, NSFC key international joint project under Grant No. 62120106007, and Interdisciplinary Research Program of HUST under Grant No. 2024JCYJ022.

secondary gateways, optical splitters, fiber links, and Wi-Fi environments, has introduced new operational challenges. First, fault localization is challenging. Issues such as slow network speed, connection interruptions, and high latency may originate from multiple sources, including abnormal optical power, hardware or software failures, configuration errors, Wi-Fi interference, or user-side problems. Actually, PON-side issues do not directly cause Wi-Fi performance degradation but rather affect the performance of the entire end-to-end link. For instance, optical power attenuation can lead to high bit error rates and packet retransmission in the link, while optical line terminal (OLT) congestion can result in increased latency and packet loss. These issues degrade the Quality of Service (QoS) for data flows. Since Wi-Fi is the last segment of the network connecting to users, problems occurring in intermediate links are sometimes erroneously attributed to Wi-Fi performance degradation. Moreover, faults often exhibit complex interdependencies across multi-level devices, rendering manual inspection inefficient and error-prone. Second, root cause analysis (RCA) is time-intensive. Traditional approaches that rely on alarms and expert knowledge struggle to quickly and accurately identify the true cause amid vast volumes of multi-source data, such as performance metrics, alarm records, configuration logs, and environmental information. This results in prolonged mean time to repair (MTTR), which negatively affects user experience and operator profitability. Third, intelligent capabilities are limited. Current operational systems lack automated and intelligent analysis functions, including root cause inference, making them ill-suited for the efficient and precise management of large-scale FTTR networks. Therefore, establishing an automated operational decision-making cycle has become an urgent and critical priority.

The accelerating pace of enterprise digital transformation, combined with the deep integration of technologies such as cloud computing, big data, and artificial intelligence, has led to a significant increase in the complexity of intelligent systems. During daily business operations and the maintenance of core systems, organizations may experience system failures or performance degradation. To ensure continuous and stable system operation, it is essential to accurately identify the root cause of issues—not only locating the fault but also understanding the underlying reasons—to develop effective preventive strategies. This process is known as RCA, which focuses on employing systematic methodologies to uncover the fundamental factors behind problems, rather than simply addressing surface-level symptoms. Currently, RCA methods can be broadly classified into two categories: data-driven and causality-driven approaches. Due to the high complexity involved in achieving complete causal inference, data-driven methods, such as association rule mining, machine learning, deep learning, intelligent agents, and knowledge graphs, have become the dominant approach in practice.

The data utilized for RCA primarily consists of three types of heterogeneous data: 1) Location-time data: This type records the physical or logical location and associated timing of fault occurrences, thereby supporting the identification of failure propagation paths. For example, in FTTR networks, this includes AP deployment topology, user mobility trajectories, and timestamps of roaming events; 2) Physical data: This category captures the physical state of the system. Wireless metrics such as the received signal strength indicator (RSSI), the signal-to-noise ratio (SNR), and air interface utilization fall into this category; 3) Log and behavioral data: This type enables in-depth causal inference and includes examples such as device kernel logs, 802.11 protocol packet captures, and user authentication records. However, the heterogeneity of multi-source data and the presence of temporal delays remain key challenges. There is an urgent need to achieve efficient feature alignment and real-time analytical capabilities.

Among data-driven methods, association rule mining aims to identify abnormal combinations of attributes to locate root causes. For example, the classification based on associations (CBA) algorithm utilizes class association rules (CARs) to identify fault causes, demonstrating high accuracy in applications such as fault diagnosis in chiller systems[4]. However, the effectiveness of this method heavily depends on the setting of thresholds for minimum support and minimum confidence. Its performance tends to be unstable across different datasets, and it is prone to combinatorial explosion in high-dimensional scenarios[5–6].

Machine learning methods are extensively employed in environments with labeled data, where they demonstrate strong performance. Neural network classifiers and k-nearest neighbors (KNN) algorithms can construct feature vectors from historical data to train supervised classification models, enabling accurate fault attribution[7]. However, this approach is heavily dependent on the availability of large quantities of high-quality labeled samples. In unsupervised or semi-supervised scenarios, the interpretability and performance of these methods are often constrained.

Deep learning methods enable automatic feature extraction and pattern recognition from high-dimensional and complex data. They demonstrate superior performance in multimodal fusion and sequence modeling, allowing the application of hybrid architectures such as CNN-BiLSTM-Attention for processing alarm sequences or leveraging Transformer-based models to integrate multi-source information for fault diagnosis[8–10]. While deep learning is highly sensitive to nonlinear relationships and latent patterns, its effectiveness in causal inference remains dependent on the comprehensiveness of feature extraction and the quality of the input data. Furthermore, model training typically demands large volumes of labeled data and significant computational resources[11–12].

In recent years, knowledge graphs and intelligent agent methods have been gradually applied to RCA. Knowledge graphs enable interpretable reasoning through structured causal networks and probabilistic models. Agents built on large language models (LLMs), such as ReAct and RCAgent, utilize interactive tool calls and dynamic reasoning loops to cope with complex and dynamic fault scenarios, significantly enhancing system interpretability and real-time responsiveness[13–14]. These methods exhibit complementary characteristics. Association rule mining is suitable for extracting highly interpretable rules. Machine learning is effective for well-labeled classification tasks. Deep learning excels at handling high-dimensional and complex patterns. Knowledge graphs and agent-based methods enhance causal reasoning and interactive adaptability, thereby improving the interpretability and dynamic performance of RCA.

This paper, based on artificial intelligence methods, addresses the challenge of Wi-Fi quality degradation in FTTR networks caused by complex factors. The main contributions are as follows:

1) To tackle the difficulty of locating Wi-Fi-side quality degradation and performing RCA in FTTR networks, a multi-task learning model based on Transformer mechanisms is proposed, enabling simultaneous AP localization and root cause type identification.

2) A discrete-event simulation platform using MATLAB is designed and implemented to simulate various network degradation scenarios (e.g., weak coverage, congestion, and interference), generating a multi-label dataset for model training and validation.

3) Experimental results show that the proposed Transformer-based multi-task learning model achieves an accuracy of 96.75% in root cause classification tasks, significantly outperforming baseline models such as LSTM, GRU, Random Forest, and XGBoost. This demonstrates its superiority in temporal feature extraction and complex pattern recognition, highlighting its high practical value.

The rest of this paper is organized as follows. Section 2 introduces the system model, including the centralized/cloud wireless-optical access network (C-WAN)-based FTTR architecture, the RCA framework for performance degradation, and the channel access mechanism. Section 3 elaborates on the Transformer-based multi-task RCA algorithm, including problem modeling, model structure, and loss function design. Section 4 validates the effectiveness of the proposed method through simulation experiments, including dataset generation, baseline comparisons, and ablation studies. Future research directions are outlined in Section 5. Finally, Section 6 concludes the paper.

2 System Model

This section introduces the FTTR system model based on the C-WAN, including its architecture, the framework for RCA of performance degradation, and the channel access mechanism.

2.1 FTTR Network Architecture Based on C-WAN

As shown in Fig. 1, the C-WAN system is functionally divided into three planes: the management plane, the control plane, and the data plane. The management plane is responsible for monitoring device status, maintaining topology infor-

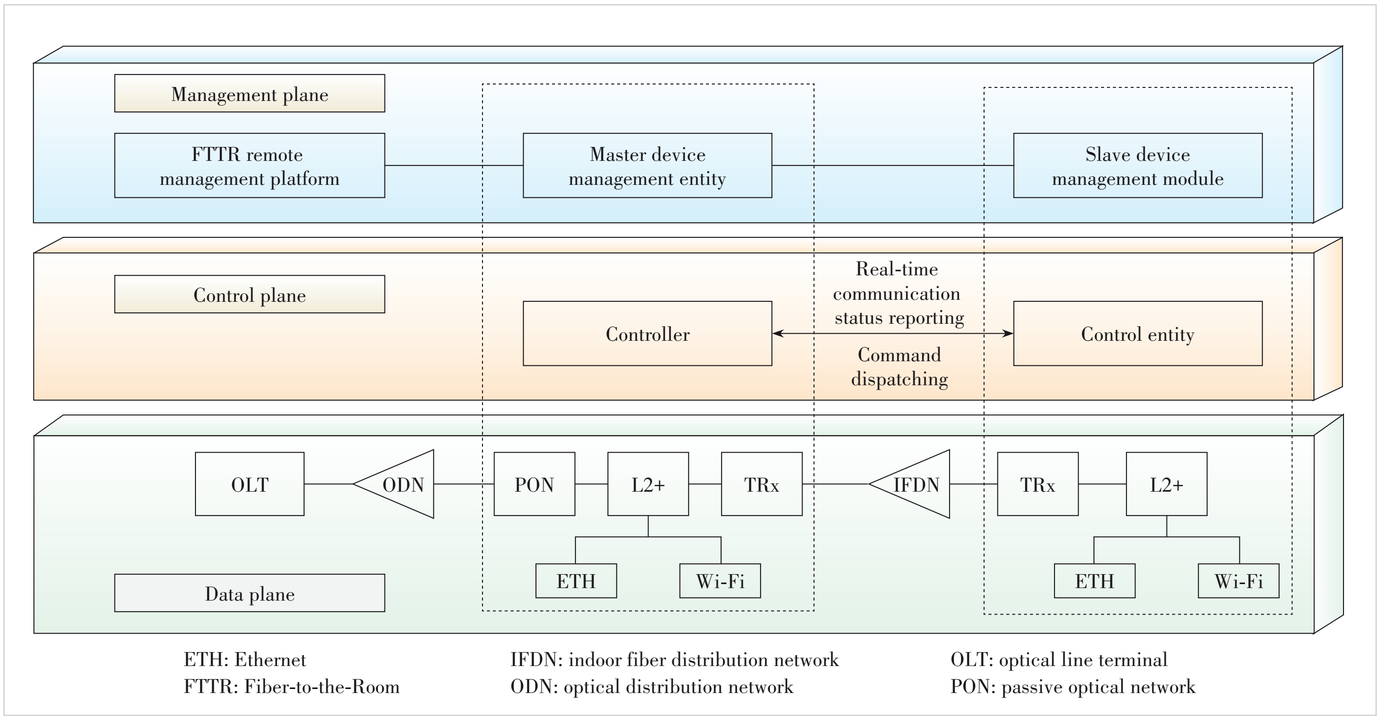

Figure 1. Centralized/cloud wireless-optical access network architecture

mation, and interfacing with operator platforms to achieve operational and maintenance objectives. The control plane primarily collects information from both air interfaces and optical links, enabling intelligent allocation and decision-making for optical-wireless resources through unified scheduling. The data plane consists of multiple devices responsible for the actual forwarding of user packets and provides transmission channels for operation and control signaling. Within this architecture, the OLT can gather various types of status information from all APs and possesses strong computational capabilities. This makes it well-suited for centralized analysis of network quality degradation causes. Detailed explanations will be provided in subsequent sections.

2.2 Framework for Root Cause Analysis of FTTR Performance Degradation

In FTTR networks where dense user and AP coexistence is common, the causes of performance degradation are complex and involve factors such as air interface collisions, channel contention, and uneven resource allocation. Therefore, this paper establishes an RCA framework for FTTR performance degradation based on the C-WAN architecture. This framework leverages real-time status data collected by the OLT, including channel utilization, collision counts, packet error rates, modulation and coding scheme (MCS), etc., and the low-latency optical fiber links to identify key factors contributing to network quality deterioration through feature extraction and correlation analysis. The analysis results can be used to guide resource reconfiguration and transmission strategy adjustments, thereby forming a closed-loop operational mechanism of "detection-analysis-optimization". This enhances network reliability and service quality.

The system model consists of a centralized OLT connected to multiple APs via optical fibers. Each AP, in turn, establishes wireless local area network (WLAN) connections with multiple stations (STAs). Leveraging low-latency optical links, the centralized OLT can promptly collect and aggregate multi-source status data generated by all APs and STAs across the network. This enables the construction of a multi-dimensional performance indicator system spanning the physical layer, link layer, and network layer. Furthermore, thanks to the edge computing module deployed at the OLT, AI models can be deployed to extract and integrate high-dimensional features. This capability allows for the identification of potential root causes of degradation, such as roaming issues, interference, collisions, congestion, and weak coverage. Based on the analysis results, optimization strategies can be formulated and delivered to the corresponding APs for execution via the C-WAN control mechanism. This architecture fully leverages the centralized advantages of C-WAN, significantly enhancing the service assurance capabilities of FTTR networks in high-density environments.

2.3 FTTR Network Channel Access Mechanism

In FTTR systems, the channel access mechanism comprises both PON and Wi-Fi components, aiming to achieve efficient and low-collision data transmission.

On the PON side, the data transmission mechanism includes downlink broadcast and uplink time division multiple access (TDMA). For downlink transmission, the OLT sends data via broadcasting, and each optical network unit (ONU) filters and receives data based on its logical identifier. For uplink transmission, multiple ONUs share the same fiber channel, and the OLT centrally schedules time slot allocations to prevent collisions. The OLT uses Gate control frames to assign transmission windows to ONUs, which then send data and Report frames within designated time slots to report their buffer status. The transmission time slots have a minimum granularity to ensure sufficient processing time for ONUs. For example, in EPON, the basic unit is 16 ns, and the minimum allocation unit is 1 024 such basic time slots (i.e., 16.384 μs). Additionally, the system periodically executes discovery and registration procedures to maintain the active status of ONUs and prevent watchdog timer timeout. Control frames (such as Gate and Report) are assigned a higher transmission priority than data frames.

In Wi-Fi networks, the enhanced distributed channel access (EDCA) mechanism is employed to enable service differentiation and QoS management. This mechanism defines four access categories (ACs): AC_VO, AC_VI, AC_BE, and AC_BK. Each AC corresponds to specific contention parameters, including minimum contention window (CWmin), maximum contention window (CWmax), arbitration inter-frame space number (AIFSN), and transmission opportunity limit (TXOPLimit). The AIFS duration is calculated as follows:

$$\mathrm{AIFS}[\mathrm{AC}] = \mathrm{SIFSTime} + \mathrm{AIFSN}[\mathrm{AC}] \times \mathrm{SlotTime} \qquad (1).$$

In the EDCA mechanism, SIFSTime and SlotTime are physical layer parameters. Each AC independently executes a backoff procedure. When the channel remains idle for the duration of the AIFS, a random backoff timer is triggered. If an internal collision occurs, the higher-priority AC gains transmission rights, while the lower-priority AC must double its backoff window. Once the transmission is successful, the contention window is reset to the CWmin value. If the transmission fails, the window continues to double until CWmax is reached.

3 Transformer-Based Multi-Task Root Cause Analysis Algorithm

3.1 Problem Formulation

The overall network architecture of FTTR comprises two major components: the PON and Wi-Fi access. In practical deployments, the PON section utilizes optical fiber as the transmission medium, offering high bandwidth, low latency, strong anti-interference capabilities, and generally stable operation with a low probability of network quality degradation. In contrast, the Wi-Fi section operates in open wireless frequency bands, making it highly susceptible to environmental interference, diverse terminal behaviors, physical obstructions, and other factors. Consequently, Wi-Fi has become the primary source of overall network quality degradation. Relevant statistics indicate that over 80% of user experience issues originate from the Wi-Fi side. Building on this, this paper focuses on analyzing and modeling Wi-Fi-related network quality issues in FTTR systems.

The analysis of FTTR network quality degradation can be formulated as a multi-task time series classification problem. Consider a network system comprising n APs. At each time step t, the system monitors N_f feature metrics, forming a feature vector $\boldsymbol{x}_t \in R^{N_f}$.

Given an observation time window of length L, the model input can be represented as a feature matrix:

$$\boldsymbol{X} = \left[\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_L\right]^{\mathrm{T}} \in R^{L \times N_f} \quad (2),$$

where each feature vector $\boldsymbol{x}_t$, $t \in [1, \cdots, L]$, includes three types of metrics: 1) AP-specific metrics: $\phi^{(j)}_{\mathrm{AP}_i}$, where $i \in \{1,2,\cdots,n\}$ denotes the AP index, and j represents the metric type (e.g., number of STAs, channel status); 2) AP performance metrics: $\psi^{(k)}_{\mathrm{AP}_i}$, reflecting the performance of each AP (e.g., throughput, interference level); 3) Global network metrics: $\omega^{(m)}$, describing the overall network state (e.g., total user count, average channel utilization).

The model outputs two vectors, Root Cause Localization (a root cause probability vector $\boldsymbol{Y}_{\mathrm{AP}} \in [0, 1]^n$ indicating the probability that each AP is the root cause AP) and Root Cause Type (a root cause type probability distribution $\boldsymbol{Y}_{\mathrm{Type}} \in [0, 1]^C$, where C is the number of root cause types). The root cause types considered in this paper include Normal, Weak Coverage, Congestion, Collision, Roaming, and Interference.

The model needs to learn a mapping function F as expressed in Eq. (3):

$$F(X; \Theta) = \left[\boldsymbol{Y}_{\mathrm{AP}}, \boldsymbol{Y}_{\mathrm{Type}}\right] \quad (3),$$

where Θ represents the parameters of the function, consisting of the Transformer-based multi-task learning model and its output heads.

3.2 Model Architecture: Transformer-Based Multi-Task Learning Model

The core architecture of the model is illustrated in Fig. 2. Each feature vector is mapped from dimension N_f to dimension N_d through a fully connected layer. Here, $\boldsymbol{W}_p$ is the weight matrix, and $\boldsymbol{b}_p$ is the bias vector.

$$\tilde{H}^{(0)} = \boldsymbol{X} \cdot \boldsymbol{W}_p + \boldsymbol{b}_p \quad (4),$$

$$H^{(0)} = \tilde{H}^{(0)} + P \quad (5).$$

Since the Transformer mechanism does not employ recurrent or convolutional structures, it cannot inherently capture the sequential relationships among elements in the sequence. The positional encoding P, a tensor with the same shape as $H^{(0)}$, is learned and added to $H^{(0)}$ to enable the model to utilize the temporal order information of the time steps in the sequence.

The Transformer encoder consists of N_L identical layers stacked together, each containing a multi-head self-attention mechanism and a feed-forward neural network. The multi-head self-attention mechanism allows any time step in the sequence to directly attend to all other time steps, effectively

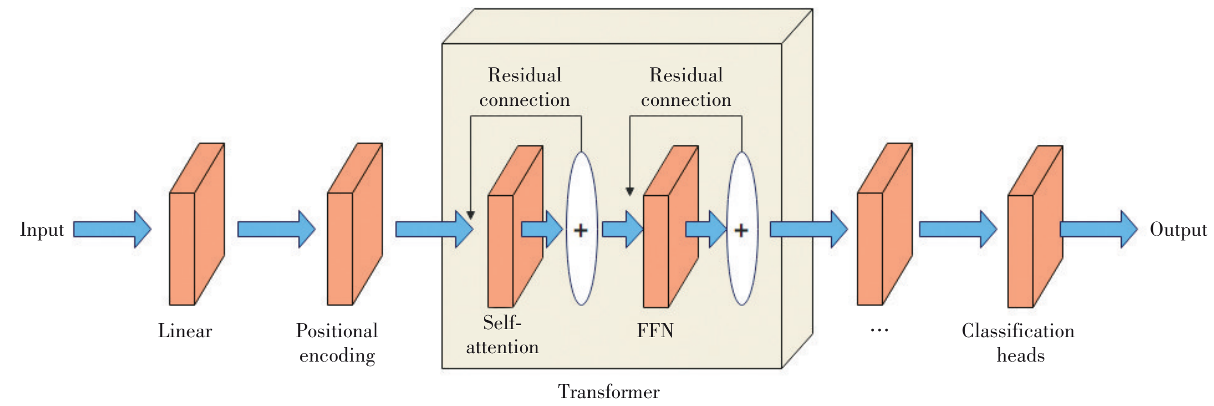

Figure 2. Multi-task Transformer mechanism

capturing long-term dependencies and interactions among features. This is crucial for analyzing the temporal causal relationships in network state metrics. The feed-forward neural network primarily applies non-linear transformations to the outputs of the self-attention layer, enhancing the model's expressive power. Each sub-layer incorporates residual connections and layer normalization, which help mitigate gradient vanishing issues in deep networks and accelerate training.

$$H^{(N_L)} = \text{TransformerEncoder}\left(H^{(0)}\right) \tag{6}$$

After processing through N_L layers of the Transformer encoder, we obtain a sequence $H^{(N_L)}$ rich in contextual information. The output of the last time step, h_{final}, is used as a comprehensive representation of the entire input sequence. It is fed into two separate classification heads to produce the model outputs Y_{AP} and Y_{Type}.

$$Y_{\text{AP}} = \text{softmax}\left(h_{\text{final}} W_{\text{AP}} + b_{\text{AP}}\right) \tag{7}$$

$$Y_{\text{Type}} = \text{softmax}\left(h_{\text{final}} W_{\text{Type}} + b_{\text{Type}}\right) \tag{8}$$

3.3 Loss Function

The total loss function is defined as the weighted sum of the two task-specific losses:

$$L_{\text{total}} = \alpha L_{\text{AP}} + \beta L_{\text{Type}} \tag{9}$$

where α and β are the weights for the two tasks. Both losses employ the cross-entropy function. The objective of model training is to minimize this total loss function.

3.4 Computational Cost Analysis

The computational cost of the model primarily originates from the initial projection layer, the Transformer encoder, and the output layer. The following analysis is based on computational complexity, where the sequence length is L, the model dimension is N_d, the input feature dimension is N_f, and the number of Transformer layers is N_L.

The operation of the initial projection layer involves matrix multiplication $\boldsymbol{X} \cdot \boldsymbol{W}_p$ and bias addition. Its complexity is $O(L \cdot N_f \cdot N_d)$. Since N_f and N_d are fixed dimensions, the complexity is linearly related to the sequence length L.

The complexity of each Transformer encoder layer is determined by the multi-head self-attention and the feedforward neural network.

1) Multi-head self-attention: The complexity of computing the attention matrix $\boldsymbol{QK}^{\mathrm{T}}$ is $O(L^2 \cdot d_k)$, where $d_k = N_d/h$ (with h being the number of heads). Since there are h heads, the total complexity is $O(h \cdot L^2 \cdot d_k) = O(L^2 \cdot N_d)$, because $d_k = N_d/h$.

2) Feedforward neural network: The complexity of the two linear transformations is $O(L \cdot N_d \cdot d_{ff})$. Since $d_{ff} \propto N_d$, the complexity is $O(L \cdot {N_d}^2)$.

Thus, the total complexity of each Transformer layer is $O(L^2 \cdot N_d + L \cdot {N_d}^2)$. The entire encoder has N_L layers, so the total complexity is $O(N_L \cdot L^2 \cdot N_d + N_L \cdot L \cdot {N_d}^2)$.

The operations of the output layer involve matrix multiplication and softmax for two classification heads. The complexity is $O(C_{\text{AP}} \cdot N_d + C_{\text{Type}} \cdot N_d)$. Since C_{AP} and C_{Type} are small and constant, this part of the complexity is negligible.

In conclusion, the total computational cost is:

$$O(L \cdot N_f \cdot N_d + N_L \cdot L^2 \cdot N_d + N_L \cdot L \cdot {N_d}^2 + C_{AP} \cdot N_d + C_{\text{Type}} \cdot N_d) \tag{10}$$

Based on the above analysis, the total computational cost of the model is primarily dominated by the Transformer encoder, i.e., $O(N_L \cdot L^2 \cdot N_d + N_L \cdot L \cdot {N_d}^2)$. The computational cost mainly depends on the sequence length L and the model dimension N_d.

4 Simulation

4.1 Dataset Generation

To train and evaluate an AI model capable of accurately determining the root causes of WLAN network issues, we developed a discrete-event simulation platform based on MATLAB. This platform simulates key mechanisms of the IEEE 802.11 protocol's MAC and physical layers to replicate various network quality degradation scenarios. The architecture of the simulator and the detailed process for generating a multi-label dataset are described below.

Our simulator follows an object-oriented design philosophy, with its core classes and interactions illustrated in Fig. 3.

The Simulator module is the core scheduling system of the simulation, responsible for maintaining the global clock and a list of APs. It advances the simulation by scanning and executing the earliest occurring events. The AP module models the behavior of real APs, maintaining a state machine that includes states such as BACK_OFF, HOLD, and SEND. The Callback module contains functions scheduled to execute at specific future time points, handling events such as the end of backoff, the start of frame transmission, and acknowledge (ACK) character timeout checks. The Events module records the details of each transmission attempt, including start time, end time, and status. The Mobile module simulates STA mobility and roaming behavior, tracking the positions of STAs and APs. It calculates the received signal strength based on real-time distance and a path loss model, triggering roaming decisions when the signal strength falls below a certain threshold. The ChannelQuality module calculates the received signal power, aggregates interference from co-channel APs and external sources, and computes the signal-to-interference-plus-noise ratio (SINR). It also determines the packet error

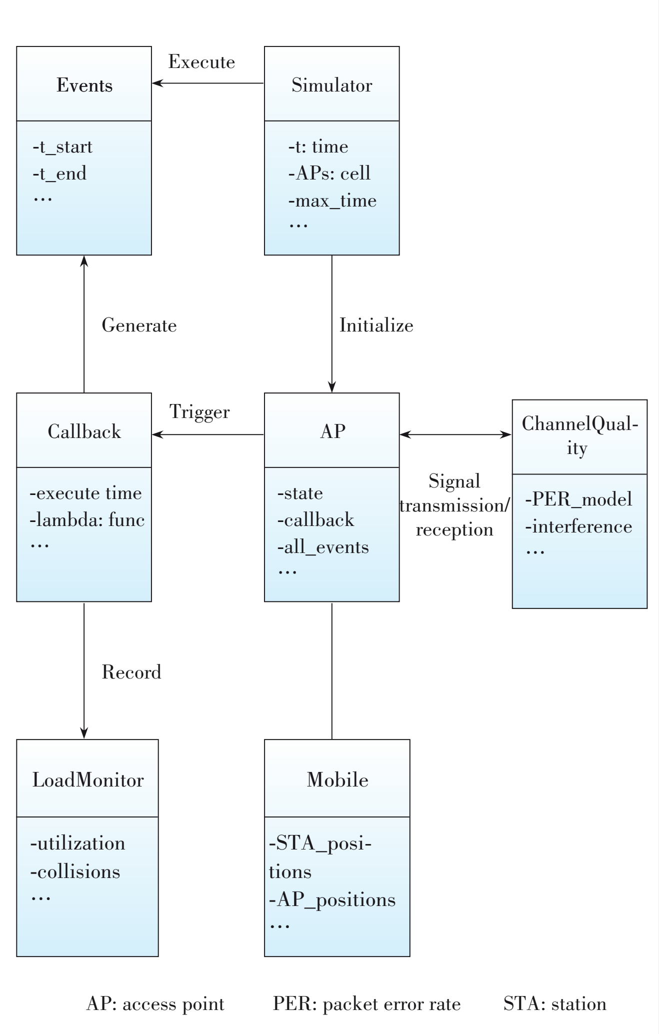

Figure 3. Simulator modules: core classes and interactions

rate (PER) using a PER-SINR mapping model, taking into account different modulation and coding scheme (MCS) levels. The LoadMonitor module monitors network load conditions, collecting statistics such as channel utilization, queue length, and collision count for each AP, and subsequently calculates the collision rate.

Data generation is achieved through multiple simulation runs, each configured with specific parameters to simulate particular problem scenarios. Roaming issues are simulated using STA mobility models, where STAs continuously move between two APs; Coverage deficiencies are simulated by setting low AP transmission power or placing STAs at the edge of coverage areas; Congestion issues are simulated by increasing the number of APs or raising the STA's transmission rate to elevate the load; Interference issues are simulated by adjusting AP topology to cause mutual interference or by introducing external APs outside the FTTR system; Severe collision scenarios are simulated by configuring small contention windows or a large number of STAs.

The dataset generation process integrates the previously mentioned modules. The first step is simulation initialization, where parameters such as the number of APs, number of STAs, physical locations, transmission power, and traffic load are configured. All AP, Mobile, ChannelQuality, and LoadMonitor modules are initialized.

During the simulation, the simulator employs an event-driven time-step advancement mechanism. At each sampling interval, it performs unified sampling of all APs to create time-stamped data samples. Each sample includes the following feature sets:

1) Network load features: the total number of STAs, the number of STAs currently associated with each AP, and system load imbalance degree (used to quantify load distribution differences among APs);

2) Channel state features: channel number occupied by each AP, the number of co-channel interfering APs, SINR for each AP, and RSSI for each AP;

3) Performance metrics: total downlink throughput per AP, maximum communication delay, frame retransmission rate, channel utilization, and the MAC-layer data frame collision rate;

4) Roaming behavior features: the number of roaming event triggers and that of failed roaming attempts per AP.

Each sample is automatically annotated with two labels. The root cause label identifies the primary network problem. The categories include roaming issues, coverage deficiency, congestion, interference, severe collisions, and normal conditions. The labeling logic is determined by the simulation configuration. The problem AP label identifies the device ID of the AP responsible for the root cause. For coverage deficiency and roaming issues, this label is the ID of the currently serving AP; for interference issues, it is the ID of the primary interfering AP; for congestion and severe collision scenarios, it is the ID of the overloaded AP. Finally, the feature vectors of all samples and their corresponding labels are stored in a matrix form to create the final dataset for model training.

In the subsequent experiment, using the aforementioned method, we conducted simulations for six different scenarios, each simulated 300 times with a time-step of 300 s. By changing the random number seeds, a dataset comprising 540 000 samples with 39 features was constructed. The dataset was partitioned into training, validation, and test sets in a 70%, 15%, and 15% ratio, respectively. Both the root cause label and problem AP label exhibit a uniform distribution.

4.2 Baseline Model Performance Comparison

The dataset generated using the aforementioned method was used to train the models. As shown in Fig. 4, as the number of training iterations increases, the loss gradually decreases, and the accuracy gradually improves, eventually stabilizing.

To comprehensively evaluate the performance of the proposed Transformer-based multi-task learning root cause analy-

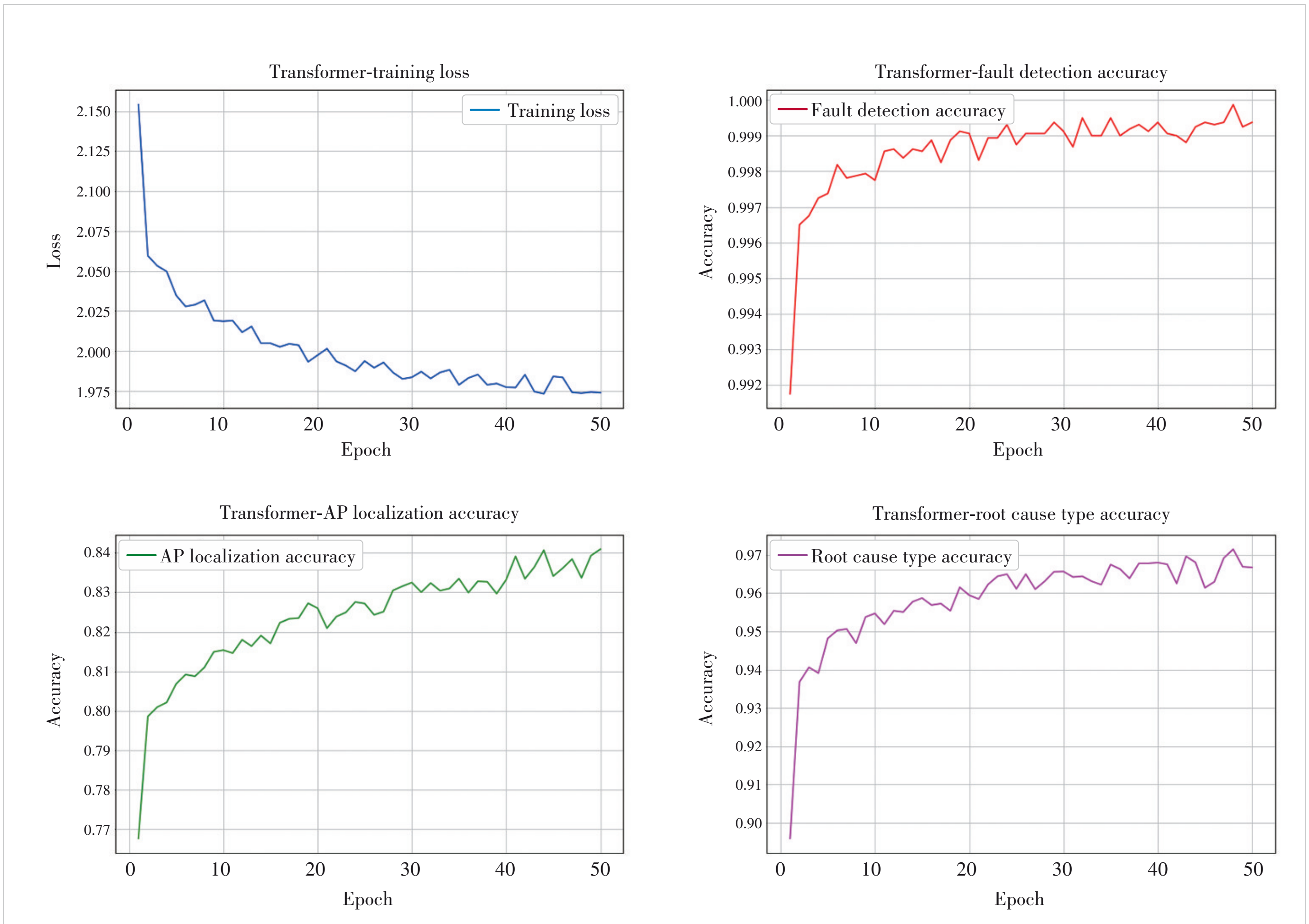

Figure 4. Baseline model performance comparison: training results

sis algorithm for FTTR networks, we compared it against several representative baseline models, including classical sequence models and traditional machine learning methods. The selected baseline models are two sequence models: Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), and two traditional machine learning models: Random Forest and eXtreme Gradient Boosting (XGBoost).

As shown in Fig. 5, the proposed Transformer-based multi-task learning model achieved the highest accuracy in both AP localization and root cause type identification tasks, demonstrating its excellent capability in temporal feature extraction and complex pattern discrimination.

•AP localization task: The Transformer-based multi-task learning model achieved an accuracy of 80.92%, outperforming LSTM (79.54%), GRU (79.02%), Random Forest (74.64%), and XGBoost (76.29%). This indicates that the self-attention mechanism is superior to recurrent structures and tree models in capturing global dependencies. However, the overall accuracy rates were relatively low, primarily because when network quality degrades, non-root-cause APs may exhibit similar features. For example, if AP1 is interfered with by AP2, AP2 may also experience interference from AP1, resulting in highly similar features between the two APs.

•Root cause type classification task: The Transformer-based multi-task learning model achieved an accuracy of 96.75%, surpassing all baseline models. Traditional machine learning methods performed significantly worse in this task, highlighting the advantage of deep sequence models in fine-grained, multi-category classification scenarios.

In summary, the Transformer-based multi-task learning model is well-suited for root cause analysis in FTTR network quality degradation scenarios.

4.3 Ablation Study on Transformer Mechanisms

To conduct an in-depth analysis of the contributions of key modules in the Transformer-based multi-task learning model, we performed an ablation study to systematically evaluate the individual effects of positional encoding and the multi-head attention

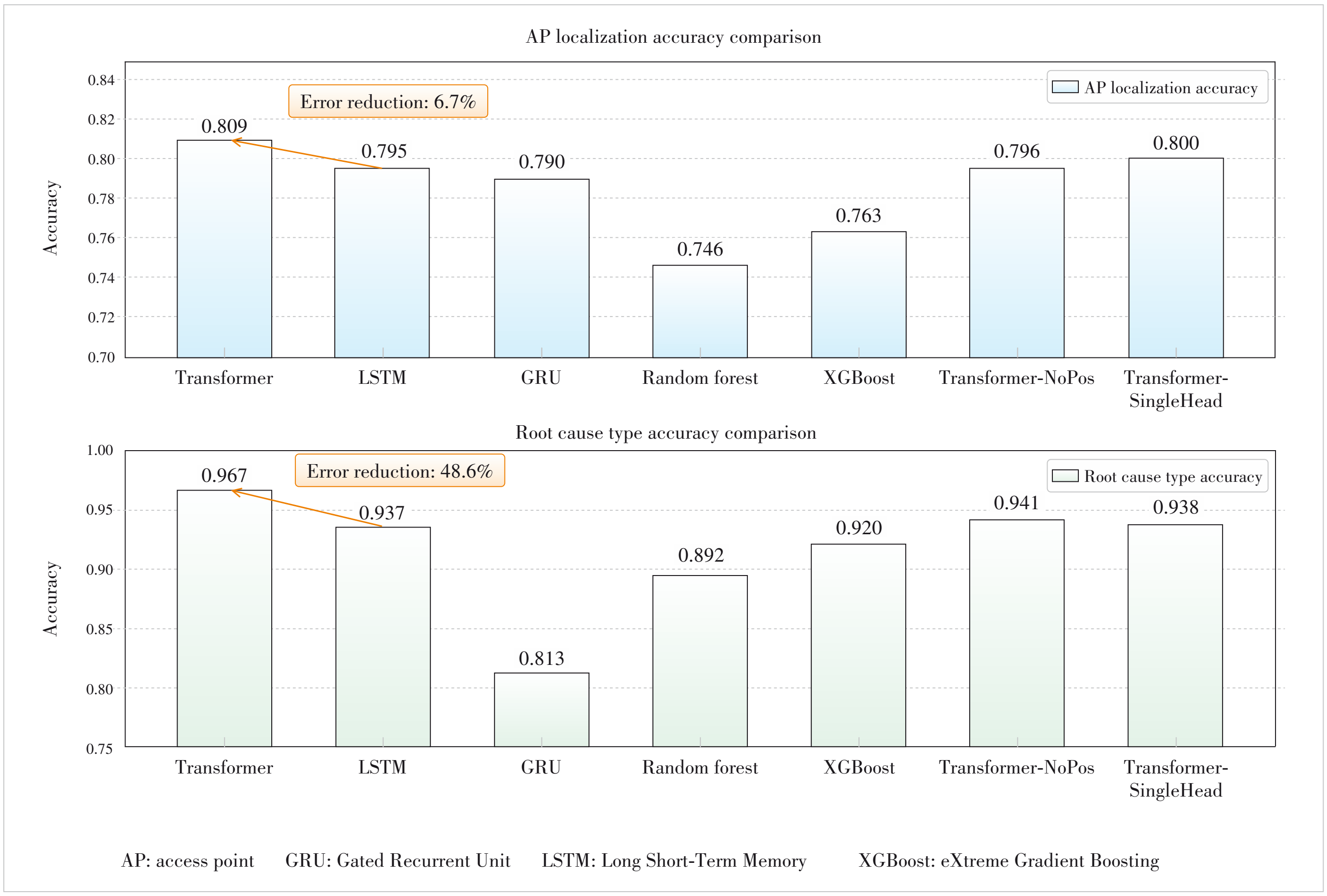

Figure 5. Accuracy comparison of AP localization and root cause type identification

mechanism. Specifically, we designed two variant models. 1) Transformer-NoPos: The positional encoding module was removed; 2) Transformer-SingleHead: The multi-head attention mechanism was replaced with a single-head attention mechanism.

Upon removing positional encoding, the AP localization accuracy dropped by 1.35%, and the root cause type classification accuracy decreased by 2.63%. These results indicate that encoding sequential information is essential for fault sequence interpretation. Without explicit positional cues, the model has difficulty distinguishing the temporal order of events. The relatively moderate performance degradation is mainly due to the use of simulation-generated data, where each scenario is preconfigured with a specific fault type. For instance, in the coverage deficiency scenario, the entire simulation consistently exhibits coverage-related issues, and the features exhibit minimal temporal variation.

Replacing multi-head attention with single-head attention led to performance declines in both tasks. The root cause classification accuracy decreased by 2.95%, confirming that the multi-head structure enhanced the model's expressive capacity and robustness by integrating diverse features from multiple representation subspaces.

Overall, the ablation study demonstrates that both positional encoding and the multi-head attention mechanism are critical components of the Transformer-based multi-task learning model, thus significantly contributing to its accuracy and generalization performance in root cause analysis tasks.

5 Future Work and Challenges

The current model faces two major challenges. First, it relies on simulated data with inherent limitations, resulting in insufficient generalization capability and credibility due to the absence of real data; second, it performs poorly in root-cause AP localization. Accordingly, we will address these two aspects separately by elaborating on the current issues and proposing future improvement plans.

Currently, our experiments rely on simulated data. There are significant differences between real data and simulated data, among which we consider the following to be particularly important: 1) The simulated data does not emulate real queues, thus lacking scenarios of overflow-induced packet loss, and the measured latency data does not include waiting delays in queues; 2) The simulation of actual service flows is inadequate, as our packet transmission process is relatively stable and lacks simulation of burst traffic; 3) Transmission control protocol (TCP) mechanisms (such as sliding windows, congestion control) are difficult to simulate. Therefore, we plan to use real data in the future to enhance the credibility of our model.

Through collaboration with the industry, we have currently set up an FTTR test environment. In the future, we will first integrate a timer within the AP to periodically measure the parameters and metrics we need, and transmit the data to a PC via Ethernet (without interfering with Wi-Fi data transmission). Subsequently, we will attempt to label the data based on real environmental data, which will be a challenging yet critical task. We cannot simply label data based on specific metrics (for example, marking it as a "coverage issue" when the RSSI falls below a certain threshold), as this would cause the model to learn our labeling methodology, which is not the desired outcome. Therefore, the rationality of the labeling approach will directly impact the final accuracy and generalization performance of our model. Finally, we will retrain and fine-tune our model using real environmental data to enhance its practical utility.

Regarding root-cause AP localization, our analysis reveals that localization performance is the poorest in congestion scenarios. Further data investigation indicates that this is due to limitations inherent in the simulated data itself. Specifically, the simulation fails to accurately model TCP-related mechanisms and does not account for queuing delays caused by Wi-Fi queues. As a result, the round trip time (RTT), which should ideally reflect feature differences between root-cause and non-root-cause APs in congestion scenarios, fails to exhibit discriminative characteristics. In addition to utilizing real-world data, we will explore the following approaches in future work to further enhance model performance in root-cause AP localization tasks: 1) incorporating device-specific attention mechanisms, 2) introducing additional highly discriminative features, and 3) experimenting with models trained on single-AP data.

6 Conclusions

This paper addresses the challenges of identifying the source of quality degradation and analyzing the complex root causes in FTTR networks by proposing a multi-task root cause analysis model based on the Transformer mechanisms. The model is capable of simultaneously detecting faulty APs and determining the types of root causes. Experimental results show that the model achieves strong performance on a simulated dataset, with a root cause classification accuracy of 96.75%, surpassing several traditional machine learning and deep learning baselines. Ablation studies further confirm the critical role of positional encoding and the multi-head attention mechanism in enhancing model performance. However, during network quality degradation, non-root-cause APs may display similar symptoms, leading to relatively lower accuracy in fault localization. Additionally, the current study is based on simulated data, where temporal feature variations may not fully capture real-world dynamics. Future work will focus on validating the model's generalization using real-world network data, exploring online learning mechanisms, improving interpretability, and integrating technologies such as knowledge graphs to enhance the transparency and reliability of the reasoning process. This study offers an effective technical solution for intelligent operation and maintenance in FTTR networks, contributing to improved network reliability and user satisfaction.

References

[1] Broadband Development Alliance. Fiber-to-the-Room (FTTR) white paper [R]. Beijing, China, 2022
[2] PFEIFFER T, DOM P, BIDKAR S, et al. PON going beyond FTTH (invited tutorial) [J]. Journal of optical communications and networking, 2022, 14(1): A31 – A40. DOI: 10.1364/JOCN.439241
[3] ETSI GR F5G 001 V1.1.1. Fifth Generation Fixed Network (F5G); F5G Generation Definition [R]. Sophia Antipolis, France: ETSI, 2020
[4] LIU J Y, SHI D L, LI G N, et al. Data-driven and association rule mining-based fault diagnosis and action mechanism analysis for building chillers [J]. Energy and buildings, 2020, 216: 109957. DOI: 10.1016/j.enbuild.2020.109957
[5] CHEN W C, TSENG S S, WANG C Y. A novel manufacturing defect detection method using association rule mining techniques [J]. Expert systems with applications, 2005, 29(4): 807 – 815. DOI: 10.1016/j.eswa.2005.06.004
[6] SIM H, CHOI D, KIM C O. A data mining approach to the causal analysis of product faults in multi-stage PCB manufacturing [J]. International journal of precision engineering and manufacturing, 2014, 15(8): 1563 – 1573. DOI: 10.1007/s12541-014-0505-8
[7] MA Q P, LI H Y, THORSTENSON A. A big data-driven root cause analysis system: application of machine learning in quality problem solving [J]. Computers & industrial engineering, 2021, 160: 107580. DOI: 10.1016/j.cie.2021.107580
[8] JAVANBAKHT N, NESHASTEGARAN A, IZADI I. Alarm-based root cause analysis in industrial processes using deep learning [EB/OL]. (2022-03-21) [2025-09-16]. https://arxiv.org/abs/2203.11321
[9] LIU P, XU H W, OUYANG Q Y, et al. Unsupervised detection of microservice trace anomalies through service-level deep Bayesian networks [C]//Proceedings of IEEE 31st International Symposium on Software Reliability Engineering (ISSRE). IEEE, 2020: 48 – 58. DOI: 10.1109/ISSRE5003.2020.00014
[10] MAMPAKA M M, SUMBWANYAMBE M. Poor data throughput root cause analysis in mobile networks using deep neural network [C]//The 2nd Wireless Africa Conference (WAC). IEEE, 2019: 1 – 6. DOI: 10.1109/AFRICA.2019.8843409
[11] ZHAO M H, ZHONG S S, FU X Y, et al. Deep residual shrinkage networks for fault diagnosis [J]. IEEE transactions on industrial informatics, 2020, 16(7): 4681 – 4690. DOI: 10.1109/TII.2019.2943898
[12] CHOI M, KIM T, LEE J P, et al. An empirical study on root cause analysis and prediction of network failure using deep learning [C]//Proceedings of International Conference on Information and Communication Technology Convergence (ICTC). IEEE, 2021: 741 – 746. DOI: 10.1109/ICTC52510.2021.9621097
[13] ROY D, ZHANG X C, BHAVE R, et al. Exploring LLM-based agents for

root cause analysis [C]//The 32nd ACM International Conference on the Foundations of Software Engineering. ACM, 2024: 208 – 219. DOI: 10.1145/3663529.3663841

[14] WANG Z F, LIU Z C, ZHANG Y Y, et al. RCAgent: cloud root cause analysis by autonomous agents with tool-augmented large language models [C]//The 33rd ACM International Conference on Information and Knowledge Management. ACM, 2024: 4966 – 4974. DOI: 10.1145/3627673.3680016

Biographies

YU Weichao received his BS degree in communication engineering from the School of Electronic Information and Communications, Huazhong University of Science and Technology, China in 2024, where he is currently pursuing his MS degree. His research interests include FTTR, Wi-Fi, and artificial intelligence.

LIU Yang received his BS degree in electronics science and technology from Huazhong University of Science and Technology, China in 2018, where he is currently pursuing his PhD degree. His research interests include Wi-Fi 7 and FTTR technologies.

ZHANG Junxiong received his ME degree in electronic information engineering from Huazhong University of Science and Technology, China in 2021. He is currently pursuing his PhD degree at the same institution. His research interests include Det-Wi-Fi and FTTR technologies.

YE Junliang received his BS degree in communication engineering from China University of Geosciences in 2011, and PhD degree from Huazhong University of Science and Technology, China in 2018. His research interests include heterogeneous networks, stochastic geometry, mobility-based access models of cellular networks, millimeter wave communications, and next-generation wireless communications.

GE Xiaohu (xhge@mail.hust.edu.cn) received his PhD degree in communication and information engineering from Huazhong University of Science and Technology (HUST), China in 2003. He was a researcher at Ajou University, Republic of Korea, and Politecnico di Torino, Italy from 2004 to 2005. He has been with HUST since 2005 and is currently a full professor at the School of Electronic Information and Communications, HUST. He is also an adjunct professor with the Faculty of Engineering and Information Technology, University of Technology Sydney, Australia. His research interests include mobile communications, traffic modeling in wireless networks, green communications, and interference modeling in wireless communications. He was the recipient of the Best Paper Award at IEEE Globecom 2010. Prof. GE is the Chinese representative for the International Federation for Information Processing (IFIP). He serves as an associate editor for *IEEE Wireless Communications*, *IEEE Transactions on Vehicular Technology*, and *IEEE Access*.