

MLNETXAI-FUSION: A HYBRID INTELLIGENT ANALYTICS SELF-HEALING FRAMEWORK FOR MACHINE LEARNING BASED DIAGNOSIS AND MITIGATION OF NETWORK ISSUES IN WIRELESS NETWORKS

PEDIREDLA SRILATHA ^{1a, 1b}, Dr. SUNEEL PAPPALA ^{2*},
Dr. SRIPADA RAMA SREE ^{3*}

^{1a} Research Scholar, Department of CSE, University of Technology, Vatika, Jaipur, India.

^{1b} Assistant Professor, Department of CSE, Aditya University, Andhra Pradesh, India.

² Professor, Department of CSE, University of Technology, Vatika, Jaipur, India.

³ Professor, Department of CSE, Aditya University, Andhra Pradesh, India.

E-mail: ^{1a} srilathaut@gmail.com, ^{1b} srilatha.p@adityauniversity.in

^{2*} suneelpappala@gmail.com, ^{3*} ramasree_s@adityauniversity.in

ABSTRACT

As the number of IoT devices and wireless distributed environments increases, the wireless communication systems are more susceptible to network problems, such as congestion, adverse traffic flow, network route instability, packet loss, decreasing latency, and hacking. These conventional network monitoring tools lack adaptivity, scalability, and static methods, making it difficult to support adaptively diagnose and autonomously mitigate network issues. In this paper, the authors have come up with a new Hybrid Intelligent Analytics Self-Healing Framework for Machine Learning Based Diagnosis and Mitigation of Network Issues in Wireless Networks (MLNetXAI-Fusion) to solve the aforementioned problems. These are integrated into an intelligent architecture: Adaptive Feature Engineering, Deep Behavioural Anomaly Mining, Hybrid Ensemble Learning, Graph-Based Contextual Intelligence, Explainable Analytics, Trust Aware Optimization and Reinforcement Learning Self-Healing Strategies. The learning latent traffic behavior is achieved with deep autoencoder, and in the second step, the abnormal network activities are detected by conducting the reconstruction analysis. Finally, a hybrid XGBoost-LightGBM ensemble model is employed to effectively detect network problems and the communication relationship between the nodes and the propagation of anomalies in context are modeled using a Graph Attention Network. The SHAP and LIME modules are Explanatory Analytical Modules which are used to improve Interpretability and Trustworthy Decision Analysis. Furthermore, an adaptive self-healing engine estimates the attack level, isolates suspicious nodes, optimizes the routing behavior and optimizes the Quality of Service (QoS) parameters with the reinforcement-based routing optimization algorithm. The experimental performances have been tested with the reference test datasets of wireless network intrusion, and the result shows that the performance of the three modules: anomaly detection, trust stability, QoS optimization and autonomous recovery, is better than the conventional methods. The proposed framework gives an intelligent, wireless network management and network autonomous recovery deployment and scaling solution.

Keywords: *Hybrid Intelligent Analytics, Autonomous Network Management, Explainable Analytics, Trust-Aware Networking, Network Anomaly Detection, QoS Optimization, Wireless Networks, Self-Healing Networks, Graph Attention Network, Reinforcement Learning.*

1. INTRODUCTION

The wireless communication networks have thus been a part of the modern communication infrastructure in recent decades and have enabled applications such as the Internet of Things (IoT), edge computing, smart healthcare systems, intelligent transportation systems, industrial automation, environmental monitoring and smart cities. As the number of connected devices continues

to increase and wireless networks are constantly evolving, network congestion, route instability, unusual traffic patterns, packet loss, intrusion into the network and degradation in the Quality of Service (QoS) are becoming problems. These problems may impact the reliability of communications, network availability and the effectiveness of service that is provided using next generation wireless infrastructure.

Classical network monitoring and intrusion detection systems are rule based, and based on thresholds or signature based. These techniques can be used to identify both known attack patterns and some types of communication errors but do not work if there are changes in communication behaviour and/or traffic characteristics. Adaptive network monitoring and intelligent diagnosis have thus caused the research into machine learning-based network analytics to be encouraged.

In recent years, a ML and Analytical Study for the improvement of Network Intrusion Detection and Communication Monitoring have been made. Hamroun et al. [1] has conducted a survey of the intrusion detection technique and problem in 5G and Wi-Fi network. Singal et al. [3] and Ahmed et al. [2] explored the LightGBM based intrusion detection and gradient-boosting based network intrusion detection, respectively. So, Nugraha et al. [4] spoke of the importance of an explainable intrusion detection and Hosain et al. [5] spoke of the adaptability and reliability of communication in wireless mesh network. These studies will provide foundations for the intelligent network monitoring, traffic classification, explainability and communication management.

An important research question is still left unanswered: Even with these advances, comprehensive network diagnosis and recovery is still not possible when identifying abnormal network behaviour. A detection or classification model may be able to detect or classify suspicious traffic, even if it does not know what the suspicious traffic is, what the relationships between the entities communicating is, or what action should be taken to mitigate the suspicious traffic. Likewise, trust, the communication situation and the quality-of-service objectives should be taken into account when restoring a network and optimizing network routes. The challenges lead to an investigation of a new integrated solution, which combines anomaly detection, classification of network issues, contextual analysis, explainable solution of decisions and adaptive recovery.

It is a study that takes both these complementary capabilities into account and fills in the missing research gap in a comprehensive framework for diagnosis and mitigation of wireless networks. The focus of this research is particularly on the incorporation of deep behavioural anomaly mining, hybrid ensemble classification, graph-based contextual intelligence, explainable analytics, trust-aware optimization and self-healing using reinforcement learning. The aim is to provide a

discussion of how these can be used together to identify/prioritise/network issues, analyse and explain them and help to mitigate.

To overcome this research gap, a novel Hybrid Intelligent Analytics Self-Healing Framework for Machine Learning Based Diagnosis and Mitigation of Network Issues (MLNetXAI-Fusion), is proposed for Wireless Networks in this study. It is comprised of two deep learning models: deep Autoencoder for the behaviour anomaly mining, and hybrid model (XGBoost – LightGBM) for network issue classification. In order to encode contextual relationships between communication entities, a Graph Attention Network (GAT) is added. To aid in the interpretation of the results of the model prediction, SHAP and LIME are used and to drive the routing, trust-aware evaluation and adaptive routing with reinforcement learning are added to the proposed self-healing approach. The investigated setup and performance analysis is assessed and compared to the other setups and to an ablation study, on a component wise basis.

In this study, the following is aimed:

1. Develop an integral solution that can offer intelligent diagnosis and mitigation of network problem in wireless communication environment.
2. In this paper, the authors aim to explore the deep Autoencoder-based behavioural anomaly mining for abnormal communication patterns.
3. To evaluate the hybrid XGBoost–LightGBM classification, against the individual classifiers.
4. Adding GAT contextual analysis modelling relationships between communication entities.
5. To use SHAP and LIME to help with the understanding of model predictions.
6. To combine the evaluation and self-healing of the network recovery actions, so that trust is taken into account.
7. Discuss adaptive routing from a network stability and some relevant QoS goals perspectives with regards to reinforcement learning.
8. To validate the framework by comparing, component-based assessment and ablation study under given set of experimental conditions.

The importance of this study is that the research has been done on a combination of network

detection, diagnosis, explanation and recovery, using the same analysis model. This study will not only consider the accuracy of the classifications, but also individual components of analyses to meet the general objective of intelligent network management. So, the result should be considered with caution as the hybrid classifier might not perform best when compared with each individual classifier.

We discuss briefly the existing works on anomaly detection, hybrid ensemble classification, contextual learning on graph, explainable AI, trust-aware optimization and self-healing mechanisms and their relationship with our works. To summarize the work done in the fields of anomaly detection, hybrid ensemble (HE) classification, contextual learning with graph, explainable AI, trust-aware optimization, and self-healing, and the gap that was covered by this work, we present the related work in the following Section 2. A high-level overview of the overall flow of the architecture and proposed MLNetXAI-Fusion framework is presented in Section 3. The research methodology and the research execution protocol is explained in Section 4 that contains details of the Dataset, Pre-processing procedures, Experimental set-up and the evaluation parameters. A detailed methodology and algorithms adopted in the proposed framework is explained in section 5. The experimental results and discussion are presented in Section 6, a comparative analysis, ablation study and computational complexity and scalability analysis are presented in Section 7 respectively. Lastly in Section 8, the study results, implications and limitations are discussed and suggestions for future research are given.

2. RELATED WORK AND LITERATURE SURVEY

Machine-learning based anomaly detection, deep learning, ensemble classification, graph based contextual learning, explainable AI, trust-aware optimization and autonomous self-healing are some areas of research in intelligent wireless network management. In the previous studies different aspects of network security and communication monitoring, traffic analysis, adaptive network management have been discussed. In this section the above research directions are summarized briefly, some of the limitations mentioned in the literature are noted, and the research gap that motivates the proposed research of the MLNetXAI-Fusion framework are defined.

2.1 Machine-Learning-Based Network Anomaly Detection

Anomaly detection using machine learning techniques has been studied for the purposes of detecting abnormal communication patterns, malicious traffic and network intrusion detection systems (NIDS). In [6] the authors Muhammad et al. introduced the framework of LIME-based explainable artificial intelligence system for intrusion detection systems called L-xaids. The study integrates intrusion detection and explainability—how important it is to provide an explanation to decisions made in intrusion detection. The other studies in the literature that are based on machine-learning algorithms for intrusion detection and traffic classification are investigated in [17], [20], [24], [30] and [33].

These are some of the methods used to detect suspicious network activities. But their abilities have to be considered for the specific detection tasks, datasets, and evaluation procedures used. Contextual information around communication relationships and what is an appropriate mitigation action are not necessarily identified by the identification of abnormal traffic. These factors lead to the research of complementary approaches to diagnosis and recovery of networks.

2.2 Deep Learning and Autoencoder-Based Anomaly Detection

Deep-learning techniques to learn complex traffic representations and identify abnormal communication behaviour have been explored. In the current manuscript we talk about reconstruction-based behavioural learning in relation to the work of Bai et al. [12] and a variational-autoencoder approach credited to Watkins et al. [13]. When used in discussion of the learned traffic representations and the anomaly detection, the following studies can be used. Other works found in the literature are hybrid deep-learning-based intrusion detection methods [20], network monitoring using deep-learning [21] and feature-representation learning for intrusion analysis [25]. These methods can be a part of the understanding analysis of complex traffic patterns, but the goals of the study and model limitations differ among the studies.

The detection of anomalies does not provide explanation/diagnostic or independent recovery or communication analysis in context. As part of an integrated network diagnosis and mitigation framework, MLNetXAI-Fusion investigates the task of Autoencoder based behavioural anomaly mining.

2.3 Hybrid Ensemble Learning for Network Classification

Ensemble learning is a machine learning technique that is used for classification problems by aggregating the outputs of several predictive models. To classify structured data, two data boosting methods - XGBoost [8] and LightGBM [9] have been created. Other ensemble-based intrusion detection system presented by Zhong et al. [24] and hybrid deep learning-based intrusion analysis system presented by El-Sofany et al. [25] were also discussed in the already available manuscript. These studies are a base to comprehend individual classifiers and combined learning methods for network analytics applications. But effectiveness of an ensemble must be determined by experimental comparison of ensemble with its component models based on a common set of data with a common measure. The use of several classifiers does not necessarily mean that its accuracy or generalization will be improved. The proposed MLNetXAI-Fusion framework investigates hybrid XGBoost-LightGBM classifiers to classify issues in a network. In the experimental part, the classification results of the individual classifiers and the combined classifiers are evaluated and the performance observed at the reported operating conditions is reported based on the above results.

2.4 Graph-Based Learning for Contextual Network Intelligence

Graph-based learning can capture relationships among entities as nodes and edges and can represent contextual data in interrelated systems. Veličković et al. [7] proposed Graph Attention Networks for learning representations from graph structured data with the use of attention mechanisms. There is also a discussion of other related works such as graph-based communication analysis from [26], relational communication monitoring from [27], temporal graph learning from [29] and graph-based anomaly interpretation from [32] in the current manuscript. The following are research directions related to analyzing the relations between network entities and understanding communication behaviour in context: Their representation in a graph, their learning objectives and their reported evaluation results should be used to evaluate and compare them. Graph-based learning may be helpful but does not necessarily lead to decisions that can be explained, trust-aware optimization or self-recovery of the network. The MLNetXAI-Fusion combines the Graph Attention Networks to investigate the contextual communication intelligence in addition to

anomaly detection and anomaly classification in the network.

2.5 Explainable AI for Network Diagnosis

Explainable AI seeks to explain a machine-learning prediction and determine the features or factors that impact the model's prediction. LIME-based explainability for intrusion detection is discussed by Muhammad et al. [6] and LIME approach for the explanation of individual model predictions is introduced by Ribeiro et al. [11]. The previous manuscript also covers explainable intrusion detection (EID), discussed by Nugraha et al. [4] and explainable feature attribution (EFA), discussed by Javed et al. [10]. These studies have shown that it is critical to consider analytical decisions and not just the prediction output. An explanation, however, is not alone sufficient to ensure that a prediction is right, or that a recovery action selected for a network will work. The diagnostic task and the operational goals of the network must, therefore, be taken into account when considering explainability.

In MLNetXAI-Fusion, SHAP and LIME are integrated into the model prediction. They are described in the methodology part, with regard to the proposed framework.

2.6 Trust-Aware Optimization, Reinforcement Learning, and Self-Healing

Apart from this one, there are two other directions of intelligent communication systems, complementary to this one – intelligent direction of Trust-aware network management, Reinforcement learning and autonomous recovery. Adaptive routing has been mentioned in the manuscript, Trust-based network analysis has been referenced and the reinforcement-learning-based intrusion detection has been referenced by Cevallos et al. [18] while the prediction of the quality of service has been referenced by Al-Zoubi et al. [19]. It contains adaptive network management, edge-based traffic analysis and self-healing and routing which are based on [22] [23] and [29] respectively. The above-mentioned research directions have following problems: Network optimization, reliability of the network, Routing and network recovery. The technique(s) to be used should be piloted based on first research results for the States, Actions, Objectives and Measures of performance. It's different in terms of capabilities; it would have some different roles in the network handling as the trust evaluation, reinforcement learning and self-healing.

The self-healing concept proposed has been a fusion of the concerns of a trust-aware evaluation,

adaptive routing and reinforcement learning (RL), aided by MLNetXAI. The experimental component will be reporting on the reported estimation of such components, and the relationship with the Network Management goals.

2.7 Comparative Analysis and Identified Research Gap

These studies reviewed have considered different features of intelligent networks management such as anomaly detection, behavioral learning, explainability, trust evaluation, adaptive optimization and ensemble classification, and graph based contextual analysis. Yet, there are variations between methods since they vary, in terms of their goals and abilities. What they did and what they didn't do. They don't have to be an obstacle, but they are evaluated as one and strengths and weaknesses of these can be compared and contrasted based on the evidence they did offer in their publications.

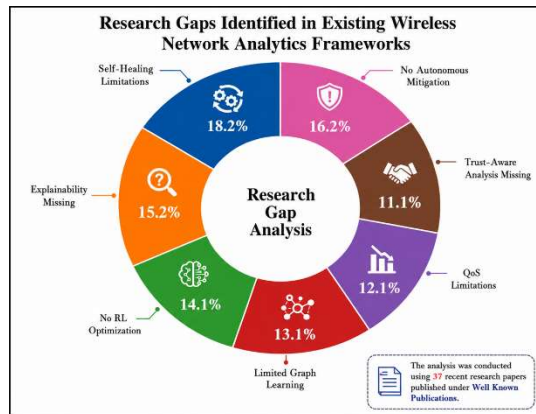


Fig. 1. Research Gaps Identified in Existing Wireless Network Analytics Frameworks

The concept of scalable intelligent communication framework of a large scale distributed wireless network is one of the research trends in recent years undertaken. From the fig. 1, we can see research gap identified from existing wireless network frameworks. This unification of complementary capabilities as the key difference has been explored in this work and led to the investigation of a single approach to diagnosis and mitigation of network problems. The proposed MLNetXAI-Fusion comprises of 4 steps namely: Autoencoder-based Behavioural Anomaly Mining, Hybrid XGBoost–LightGBM Classification, Graph Attention Network for Contextual Analysis and Explanation with SHAP/LIME and Trust aware Evaluation and Self-healing with Reinforcement Learning. The assessment of the framework will be done by the following experimental approaches and comparisons. It just describes a workflow of

analyzing steps – not that each of the analysis steps, or the whole workflow is better than what is available in the world today.

3. PROPOSED MLNETXAI – FUSION FRAMEWORK

In addition to the above, modern communication systems like Internet of Things (IoT) network, edge-assisted systems, software-defined communication environment, and distributed network architectures produce large numbers of traffic with dynamics that are susceptible to congestion, packet loss, abnormal communication, degradation of service and cyberattacks. Existing network monitoring and intrusion detection solutions (IDS) are mostly attack classification/traffic analysis based, offering no explainability or having any kind of intelligence or decision making in a trustful manner and they do not self-recover.

To overcome these limitations, this paper introduces a Hybrid Intelligent Analytics Self-Healing Framework (MLNetXAI-Fusion) to handle problem diagnosis and mitigation aspects of Wireless Network (WN) using the ML techniques. In addition to the reinforcement learning-based self-healing capabilities, it integrates all of the aforementioned features, deep behavioral anomaly mining, hybrid ensemble learning, graph based contextual intelligence, explainable analytics and trust aware optimization in a single framework.

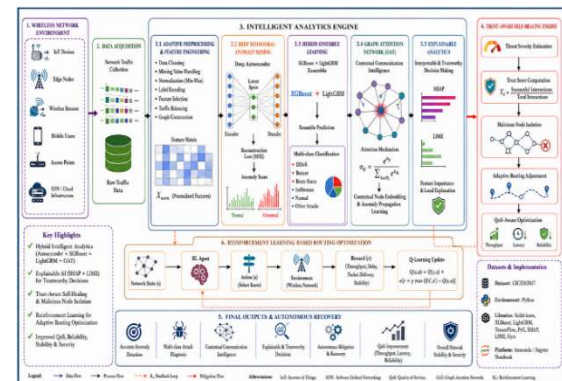


Fig. 2. Proposed MLNetXAI-Fusion Architecture for Intelligent Diagnosis, Explainable Analytics, Trust-Aware Optimization, and Autonomous Self-Healing of Network Issues

The proposed framework can have the capability to intelligently detect the abnormal behavior of the network, reason the abnormal behavior based on the context, mitigate the abnormal behavior based on trust and automatically optimize the performance of the network. The proposed

MLNetXAI-Fusion framework's architecture is displayed in figure 2.

3.1 Overall System Architecture

It can be broken down into 6 big steps towards implementing this work, namely the proposed MLNetXAI-Fusion 6 steps are:

1. A new data acquisition method, that is adaptive pre-processing is proposed.
2. Intelligent Feature Engineering
3. DBAs' name is Deep Behavioral Anomaly Mining.
4. Hybrid Ensemble Diagnosis
5. Including Contextual Intelligence + Analytics with explanation (with graphs).
6. Trust-Aware Self-Healing Optimization

Initial data will be gathered from the traffic in a network and following that the next step – pre-processing – will be performed with the aim to manage data inconsistent and normalization of the attributes of the network traffic. Finally, the parameters of the smartsys behaviour can provide information on the stability of the network, increase and/or decrease of the communication services and status of traffic in the traffic-based network and communication services reliability.

Then, reconstructed traffic will be fed back to the Deep Autoencoder for further training and reconstructed traffic will be used to attempt to model otherwise normal behaviors of the traffic and create its own latent space for otherwise normal behaviors of the traffic. Then we apply the ensemble model we get from XGBoost-LightGBM and apply the model to the new network to intelligently detect what problem is occurring in the network, and what type of attacks the network is suffering from. To model the communication relationship between network entities (e.g., traffic context network), and the relationship/dependency between nodes in the network, we use graph attention network (GAT). Explainability modules of Artificial Intelligence (XAI) are provided which would be helpful in explaining reasoning and feature attribution analysis (e.g. SHAP, LIME) to improve the reasoning based on the features considered. Lastly, based on the trust value, self-healing and optimization are performed by using reinforcement learning mechanism resulting in mitigation recommendations, suspicious node isolation, network behaviour optimization and improvement in the Quality of Service (QoS) performance.

A short explanation of the complete processing sequence as follows:
Input traffic → preprocessing → feature engineering → anomaly detection → classification → contextual

analysis → explanation → trust evaluation and recovery.

3.2 Adaptive Data Preprocessing

These traffic data sets may not be complete, include duplicate traffic information, contain gaps in their data and be dominated by particular type of traffic. The adverse effect of these incongruities is on the anomaly detection/classification performance.

In pre-processing the following are done:

1. Explicitly represent the null and inf values.
2. Duplicate traffic elimination
3. Hand coding practice of categorical attributes
4. Min-Max normalization
5. Using the technique synthetic minority oversampling technique (SMOTE) in class balancing.
6. Graph construction preparation

Normalization of the features is done by:

$$X_{\text{norm}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

Where

$X_{\min}$ – minimum value of the traffic feature

$X_{\max}$ – maximum value of traffic feature and X is value to be removed.

The purpose of preprocessing is to increase consistency of features, decrease bias and facilitate convergence of model.

3.3 Adaptive Network Intelligence (ANI) Method by Feature Engineering

It is used that changes the frequency domain of data along with number of prediction frames. A key output of the set of proposed frameworks is that it results in intelligent behavioural descriptors that are capable of describing the degree of communication reliability, network stability and the traffic abnormalities.

Engineered features are:

1. A Traffic Stability Score (TSS) is used to indicate a traffic's stability.
2. Packet Burst Ratio (PBR)
3. QOS Degradation Score (QDS): .
4. Congestion Score (CS)
5. Results have been shown based on Delay Variance Score (DVS).
6. Traffic Risk Index (TRI)
7. Adaptive Trust Score (ATS)
8. FDM = Flow Disorder Metric.

This approach is referred to as Energy Drain Estimation (EDE). The amount of variation across the traffic (or health) data is measured over a period of time. The Traffic Risk Index is the outcome of many indicators of congestion, behaviour, instability, delay variability and of traffic disorder. Similarly, Adaptive Trust Score calculates behaviour reliabilities of entities involved in communications according to interaction behavior of the network and anomalies. These not only enhance the anomaly discrimination capability, but they also better model the behaviour of these than traditional traffic features.

3.4 Hybrid Ensemble Diagnosis

Within the proposed framework (Proposed FC), there is an introduction of an anomaly mining mechanism based on Deep Autoencoder to discover the abnormal communication patterns in the files. The encoder is responsible for reducing high dimensional observations of traffic to simple representations in latent space:

$$z = f(Wx + b)$$

This is done by two sets of trainable parameters, namely weights W and bias parameters b , and by an input traffic feature vector x , and a set of latent representations z , which represent behaviours of the traffic.

The traffic pattern is recreated by the decoding unit:

$$\hat{x} = g(W'z + b')$$

Mean Squared Error (MSE) is used to calculate the reconstruction loss:

$$MSE = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2$$

The complete autoencoder model has:

Input $\rightarrow 128 \rightarrow 64 \rightarrow 32 \rightarrow 16 \rightarrow 32 \rightarrow 64 \rightarrow 128$
 $\rightarrow$ Output

The anomaly scores are obtained from the reconstruction errors. When the point with 95th percentile error is defined as dynamic threshold point, traffic samples whose error in traffic reconstruction is higher than the dynamically obtained threshold value are considered as abnormal. This module enables:

- Hidden anomaly discovery
- Latent traffic representation learning is = traffic data processing without supervision.
- Behavioural pattern mining
- Unknown attack detection
- Zero-day anomaly identification

3.5 Hybrid Ensemble Learning for Intelligent Diagnosis

An intelligent diagnostic system is implemented by making use of Hybrid Ensemble Learning Approach. The proposed framework uses two classifiers – XGBoost and LightGBM and ensemble [32] by weights and reduces the boosting of the diagnosis accuracy and generalization capability of the classification. If it's a "knowing prob" problem in the ensemble prediction then in this case there is really only one problem: can it be extrapolated out of the ensemble prediction?

$$P_{\text{fusion}} = 0.6P_{\text{XGB}} + 0.4P_{\text{LGBM}}$$

If it's PXGB model, then it is the probability of class label "PXGB" and if it's XGBoost model then it is the Probability values of class label "PLGBM".

The joint of the ensemble rain runoff is in a direction towards:

- Multi-class anomaly classification
- Detection robustness
- Generalization capability
- Feature learning efficiency
- Computational scalability

The model is able to visualize various attacks which can be performed in a multi network environment and detect "abnormal" communications behavior.

3.6 Graph Attention-Based Contextual Communication Intelligence

To summarise, there are a number of complex ways we can send a message when we are on a medium. In short: lots of complex ways exist to convey a message when on a medium. In a traditional machine learning model, the records of the network traffic are independent of each other and models tend to miss out on contextual relationships between communicating entities.

In order to overcome this drawback, the proposed framework makes use of Graph Attention Network (GAT).

Communication graph is shown as:

$$G = (V, E)$$

where V is communication nodes, and E is communication links.

The nearby nodes all have an attention coefficient of:

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_k \exp(e_{ik})}$$

$$\sum_k \exp(e_{ik})$$

where $(e_{i,j})$ is the importance of communication between node “i” and node “j”.

With GAT module, you can:

- Context-aware anomaly learning
- Communication relationship analysis
- Relational traffic intelligence

In this paper, we study the problem of anomaly propagation in the context of graphs.

- Graph-driven behavioural learning

3.7 Explainable Analytical Framework

To store the transparency and trust of intelligent diagnosis, several special modules (SHAP and LIME) exist and are integrated in the model.

What are the features that according to SHAP estimates (Shapley value analysis) are important on a global level?

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|! (|F| - |S| - 1)!}{|F|!} [f(S \cup \{i\}) - f(S)]$$

ϕ_i is a weight (contribution) of feature i to the prediction outcomes. LIME will determine a close and easier to understand "surrogate" and provide individual explanations for each of the predictions.

Explainability framework offers:

- Feature attribution analysis
- Transparent anomaly reasoning
- Local prediction interpretation

Includes the addition of feature importance visualization for the entire globe.

- Trustworthy communication intelligence

3.8 Trust-Aware Self-Healing Optimization

The embedded trust aware Engine, which is introduced by the proposed model in this paper, can:

- threat severity estimation,
- malicious node isolation,
- adaptive routing adjustment,
- QoS-aware optimization,
- communication recovery.

Trust score estimation is defined as follows:

$$T_i = \frac{\text{Successful Interactions}}{\text{Total Interactions}}$$

Low trust nodes communicate on the network; they will almost certainly not trust each other or any node in the network.

Numerous operations have been carried out to improve a self-healing action:

- communication reliability,
- autonomous mitigation,
- adaptive recovery,
- intelligent fault tolerance,
- secure communication stability.

The proposed Intelligent WDSA-Architecture (fig.2) is an adaptive and intelligent architecture, that might help in problem diagnosis and mitigation by using wireless communication, which means analyzing the received wireless signals in the sensors and taking autonomous decision on problem mitigation. It features adaptive preprocessing, powerful behavioral anomaly mining, hybrid ensemble learning, graph attention based contextual communication intelligence and explainable analytical mechanisms to ensure the trustworthy detection and decision making for behavioral anomaly. Moreover, reinforcement learning driven routing optimization and trust-based self-healing technologies are added in the architecture to enable a powerful recovery from network degradation situations. In the proposed system, periodically it observes the wireless traffic and identifies the abnormal wireless traffic and quarantine the malicious nodes, optimize the wireless traffic path to achieve better quality of service, reliability, throughput and communication stability. Moreover, mechanisms of explainable AI (XAI) like SHAP and LIME make the prediction interpretable and transparent and it can be applied in real-time to intelligent wireless network management and self-communication systems.

3.9 Reinforcing the Routing with Learning (RLAO) Method

The proposed framework uses reinforcement learning-based adaptive routing optimization to obtain the best communication recovery and routing adaptation. By reward-based optimization, the RL Agent continuously learns how to communicate through the best path and communicates with the communication environment.

The update equation to Q-learning is:

$$Q(s, a) = Q(s, a) + \alpha [r + \gamma \max_{a'} Q(s', a') - Q(s, a)]$$

where:

$Q(s,a)$ is Q-value,
 α is learning rate,
 γ is a discount factor,
 r is reward.

The reinforcement learning module learns how to optimize:

- congestion reduction,
- adaptive routing,
- throughput improvement,
- latency minimization,
- autonomous communication stabilization.

4. RESEARCH METHOD AND EXECUTION PROTOCOL

The proposed MLNetXAI-Fusion framework is used to explore the dataset, preprocessing process, model configuration and the computing environment, as well as the evaluation protocol. This section provides an explanation of the dataset, pre-processing steps, model configuration and computing environment and evaluation protocol to investigate the proposed MLNetXAI-Fusion framework. This protocol should be used descriptively, both as a description of the experiment, and as a guide to repeat the experiment.

4.1 Dataset Collection and Traffic Environment

In the proposed framework, the dataset CICIDS2017 is applied in the benchmark wireless intrusion and network traffic dataset which is one of the most popular datasets to evaluate the intelligent wireless intrusion detection and anomaly analysis systems. This data set includes realistic and malicious communications traffic, including botnet traffic, distributed denial of service (DDoS) attack traffic, abnormal communication traffic patterns, brute force attack traffic and infiltration behaviour and port scanning traffic in the context of a heterogeneous network environment.

It was decided to use the CICIDS2017 data set due to the following:

1. A realistic wireless communication behaviour (WisBeC)
2. Network attack scenarios for multi-class.
3. Large-scale traffic diversity
4. QoS degradation patterns
5. Temporal communication characteristics
6. Real World Communication suitable for Intelligent Analytics

This data contains network flow data such as:

- Flow duration
- Packet transmission rate
- Statistics of packets: Forward and Backward
- Communication frequency
- Packet inter-arrival time
- Throughput behavior
- Header information
- Routing statistics

74	Idle Mean	692703	non-null	float64
75	Idle Std	692703	non-null	float64
...				
76	Idle Max	692703	non-null	int64
77	Idle Min	692783	non-null	int64
78	Label	692783	non-null	object
dtypes: float64(28), int64(54), object(1)				
memory usage: 417.5+ MB				
None				
Missing Values:				
	Destination Port	0		
	Flow Duration	0		
	Total Fwd Packets	0		
	Total Backward Packets	0		
	Total Length of Fwd Packets	0		
	...	...		
	Idle Mean	0		
	Idle Std	0		
	Idle Max	0		
	Idle Min	0		
	Label	0		
Length: 79, dtype: int64				
Class Distribution:				
	Label	count		
	BENIGN	446001		
	DoS Hulk	231073		
	DoS GoldenEye	18293		
	DoS slowloris	5796		
	DoS Slowhttptest	5499		
	Heartbeat	11		
Name: count, dtype: int64				

Fig.3. Dataset Distribution and Communication Traffic Categories in CICIDS2017

The traffic environment which is being considered in the proposed framework are as follows:

- IoT communication systems
- Edge-assisted wireless networks
- Software-defined communication infrastructures
- Wireless communication systems that are distributed.
- Dynamic and varied communication situations (DVS).

The data set and the categories of communication traffic for CICIDS2017 data set for intelligent anomaly detection analysis are presented in Fig. 3. As shown in the figure, the data set contains also normal network traffic (BENIGN) and attacks from the various categories: DoS Hulk, GoldenEye, Slowloris, Slowhttptest and Heartbleed. It also validates that missing values were properly dealt with during the data preprocessing phase, resulting in the consistency and reliability of the data used for testing the proposed MLNetXAI-Fusion framework.

4.2 Data Cleaning and Preprocessing Strategy

Typical raw network traffic data sets include redundant communication attributes, noisy communication behavior, undefined numeric data, duplicate records, and inconsistent communication patterns. So, an adaptive preprocessing mechanism was used to achieve better feature consistence and for communication intelligence learning.

In the proposed Framework, the following pre-processing operations are used:

- The filter gets rid of null-valued traffic entries.
- Communication is an infinite value that needs to be removed.
- The duplicate communication record removal feature removes any communication records that are duplicated in the database.
- Label encoding of traffic features (categorical data).
- Min-Max normalization
- Feature balancing
- Dimensional consistency verification
- Communication graph generation

The mathematical form of the normalization process in the framework is:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

This normalization serves to stabilize the numbers, and to learn deep behavior features during the autoencoder training and classification with the ensemble.

4.3 Experimental Configuration and Algorithm Parameters

For the MLNetXAI-Fusion, the input representation is an 88-feature representation from the CICIDS2017 data set, of which 70% is used for training and 30% is used for testing. Min-max normalization and SMOTE are used for feature scaling and class imbalance, respectively, as needed. The Autoencoder is an encoder with 128, 64, 32 and 16 neurons, and a decoder with 128 (mirror), 64, 64 and 32 neurons. It is optimized with Adam optimization and Mean Squared Error (MSE) for training, and it's reported that it's trained for 50 epochs and 256 batch size. Error in reconstruction at the 95th percentile of the reconstruction errors is calculated as anomalies. In traffic classification, XGBoost and LightGBM are used. For contextual analysis, explainability & mitigation (if applicable), GAT, SHAP/LIME are used.

4.4 Hardware and Software Environment

The experiments are written in Python 3.x and Jupyter Notebook, and will use Scikit-learn, XGBoost, LightGBM, TensorFlow/Keras, PyTorch Geometric, SHAP and LIME as needed, depending on the experiment. The OS, processor, RAM, GPU and software-library versions are indicated based on the actual experimental environment in order to facilitate re-production.

4.5 Training, Testing, and Evaluation Protocol

The data is pre-processed and then divided into 70% training, and 30% testing. The data is pre-processed using transformations and the classes are balanced to prevent data leakage. To avoid data leakage, data pre-processing transformations and class balancing operations are carried out on the training data. The reconstruction errors using the Autoencoder are used to identify abnormal traffic and the preprocessed features are used to classify the traffic using the classifiers XGBoost and LightGBM. When applicable, the accuracy, precision, recall, F1 Score and ROC-AUC metrics are used to evaluate the implemented components of the fusion and supporting framework, while the QoS and trust metrics are used to evaluate the implemented supporting framework. Comparative analysis is carried out using consistent test set and evaluation procedure.

5. DETAILED METHODOLOGY AND ALGORITHM DESIGN

An efficient wireless communication analytics system needs efficient data sets preparation capability, adaptive feature engineering capability of the graph structured data, strong capability of representation learning of the anomalies and good experimental validation capability. The proposed frameworks consisting of 2 different experiments were conducted in dynamic wireless communication environment and a full experimental pipeline (MLNetXAI-Fusion) was envisioned to test the proposed framework. In Experiment 1, the performance of the anomaly detection and the contextual communication intelligence was tested, while in Experiment 2, the explainable analytical performance of the framework, the trust-aware optimization and the self-healing using reinforcement learning feature of the framework was tested. Designed in python and using the following Machine Learning libraries: Scikit learn, Tensorflow, SHAP and LIME along with a few other libraries which were documented/visualised using the Jupyter Notebook.

5.1 Adaptive Feature Engineering

The key concept of the proposed framework is to choose the most intelligent way to discriminate/communicate the context with the intelligent adaptive feature engineering. The framework is based on communication intelligence features that can be extracted from the raw communication traffic, and is representative of the dynamic behaviour of the communication network.

The features which are extracted are: Adaptive features:

- Packet flow density
- Communication frequency
- Traffic burst behavior
- Packet delivery consistency
- Throughput variation
- Latency variation
- Trust interaction score
- Routing stability
- Communication node activity
- Temporal packet distribution

With the help of feature engineering mechanism, it is improved as:

- Abnormal traffic discrimination
- Latent Communication Behaviour is one of the types of learning.
- Graph contextual intelligence
- Trust-aware communication analysis
- Stability (with regards to reinforcement learning / optimisation).

5.2 Deep Autoencoder-Based Behavioural Anomaly Mining

The proposed MLNetXAI-Fusion framework consists of a preprocessing stage for the network traffic, by applying adaptive pre-processing and intelligent behavioural feature engineering, then a deep autoencoder (AE) to extract some of the hidden behavioural representations from a high dimensional network traffic. The deep autoencoder is an unsupervised representation learning problem which can be used to produce an unknown traffic pattern detection solution for normal and abnormal traffic, while the supervised classifiers can be trained with the help of traffic instances annotated with the traffic patterns. This behavioural learning is performed at the initial stage prior to the hybrid ensemble classification and the whole system is robust against this behavioural learning due to extracting compact feature representation.

The principle of the autoencoder is very straightforward: to learn the compressed representation to be used to reconstruct the normal behaviour of the communication process with very low reconstruction error. The reconstruction errors during the learning process are relatively low for a benign communication, as the network can learn multiple behavioural characteristics of it routinely. If there is an abnormal communication flow, however, the reconstruction error will be relatively large since the reconstructed communication flows will not be similar to the ones learned. Thus, the reconstruction

error can be used as a good indicator of behaviour in anomaly mining.

Here, the designer (also known as the "encoder") encodes the designed feature vector in an engineered manner to a low dimensional latent representation.

$$\mathbf{z} = f_{\text{enc}}(\mathbf{x})$$

where $\mathbf{x}$ is the engineered behavioural feature vector and $\mathbf{z}$ is the latent behavioural embedding learned by the encoder.

The decoder recovers the original feature representation as:

$$\hat{\mathbf{x}} = f_{\text{dec}}(\mathbf{z})$$

Where $\hat{\mathbf{x}}$ is the reconstructed communication feature vector.

The reconstruction loss employed during training is the Mean Squared Error (MSE):

$$L_{AE} = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2$$

Number of behavioural features, determines values of N .

The proposed framework does not use a predefined threshold, instead adaptively learns the threshold from the distribution of reconstruction errors. This anomaly threshold is chosen at 95th percentile of the reconstruction error of training samples thus reducing the manual handling and adjustability of the system to different communication context. If the communication is done with reconstruction errors above this adaptive threshold, then the reconstruction error is considered "behaviourally anomalous" and sent to the next intelligent classification stage for detailed diagnosis. The proposed deep autoencoder is a multi-layered fully connected encoding-decoding network, employing a nonlinear and complex activation function for learning complex relations between the features of communication. The output of the bottleneck is a compressed behavioural description which preserves relevant information in the communication without discarding irrelevant information. It is an excellent representation to enhance the discriminative power of a follow-up hybrid ensemble classifier.

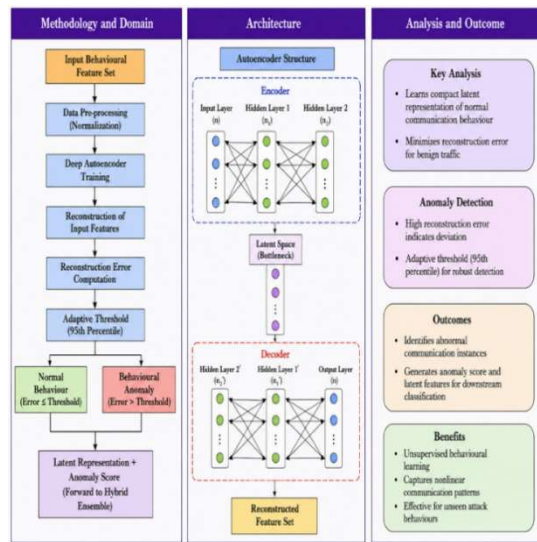


Fig. 4. Architecture of the Deep Autoencoder for Behavioural Anomaly Mining

The deep autoencoder is integrated into MLNetXAI 3, providing it with two major benefits. The first is that it allows for training the behaviour without the need for labels. Secondly, the latent representation learnt will reduce the feature redundancy and boost the effectiveness of the computation while classification. Third, the adaptive reconstruction-error analysis is another behavioural measure of confidence besides the explainability and trust analysis modules in the following phases of the framework. Algorithm 1 outlines the whole anomaly mining process within the proposed framework and Fig. 4 depicts the behavioural learning architecture.

Algorithm 1: Deep Autoencoder based Behavioural Anomaly Mining

Input:

- Network traffic dataset D that has been pre-processed.
- Created behavioural functions F

Output:

1. Latent behavioural representation
2. Reconstruction error
3. Behavioural anomaly score
4. Load engineered behaviours.
5. Normalize feature vectors.
6. Build Encoder/Decoder networks.
7. Use normal communication patterns to train the autoencoder.
8. Create latent feature representations.
9. Reconstruct the values of the feature vectors.
10. Calculate reconstruction error: Mean Squared Error.
11. Estimate adaptive anomaly threshold from the reconstruction error distribution.
12. Behaviours where communication flows are

over the threshold are marked as anomalies.

13. Pass the forward latent representations and anomaly scores to hybrid ensemble classifier.

The behaviour intelligence layer of proposed MLNetXAI-Fusion framework are Anomaly scores while latent behavioural representations are the anomaly scores obtained from the proposed deep autoencoder. This can't directly classify specific number of network attack types in the proposed method we are going to do this unsupervised learning process and adaptively reconstruct error analysis. These learned latent representations, in combination with the pre-designed behaviour descriptors then enter a supervised hybrid ensemble learning module which is then used for the intelligent analysis of the multi-class network. The proposed framework could possibly leverage on the capability of behavioural anomaly and discrimination classification, both of which would help improve the accuracy of proposed framework.

5.3 Hybrid Ensemble Classification for Intelligent Network Diagnosis

The deep autoencoder learns unsupervised representation of the network's communication behaviour, and is able to detect abnormal communication activity without any fine-grained classification of the communication activities in the network. Latent feature space gives rise to communication anomalies that can have identical behavioural attributes for a variety of attack scenarios; this makes it difficult to classify these anomalies based solely on reconstruction error. In this proposed MLNetXAI-Fusion framework, the complementary learning capability of two models, namely Extreme Gradient Boosting (XGBoost) and Light Gradient Boosting Machine (LightGBM) are fused to get powerful and precise network diagnosis. The hybrid ensemble classifiers are based on the latent representations in addition to all the behavioural descriptors engineered by the deep autoencoder described in the previous section that yield anomaly-aware representations. The hidden behavioural properties are used to complement the statistical communication properties to achieve a better discrimination of benign traffic and multiple types of network attacks.

This proposed hybrid approach adopts two classifiers that are trained by gradient boosting, however, with different methods of building trees as compared to the traditional single model classifiers. While XGBoost is a sequential learning algorithm, it is powerful and scales well; it may converge faster than LightGBM that optimizes on the histogram, but may not work well as a regularization that nicely

works for XGBoost. The ultimate goal of these models is to improve stability of the classification, prevent overfitting and boost the generalization ability of these models in a heterogeneous communication setting.

The whole process of classification of hybrid is carried out in three steps:

1. Independent learning using XGBoost Classifier.
2. Self-learning using Classifier: LightGBM.
3. The so called “adaptively weighted fusion” is employed to get the final diagnosis of the network.

The idea of the proposed hybrid classification layout is shown in Fig. 5.

5.3.1 XGBoost-Based Intelligent Classification

The first part of the hybrid ensemble is built using the engineered behavioural features and deep autoencoder latent representations which are fed to an Extreme Gradient Boosting (XGBoost) supervised learning machine. The Gradient Boosting method is an ensemble of decision trees generated by Gradient Boosting: Every decision tree is created on the “residuals” of the previous boosting a try to reduce the loss of the residuals. The iterative optimization approach enables the classifier to learn very non-linear communication features to the categories of network attacks.

XGBoost has regularization (except for the second order gradient optimization) to prevent model overfitting. Therefore, training the classifier to better analyse this mixture of heterogeneous traffic (both normal and malicious) and to make it more robust are possible.

XGBoost: Classifiers will give out the prediction in the following format:

$$\hat{y}_{XGB} = \sum_{k=1}^K f_k(x)$$

Where $f_k(x)$ is represented as the decision function of the k^{th} boosted tree and K are the number of trees and number of people.

The autoencoder is trained on the knowledge of the anomalies and the embeddings (also known as features) extracted by the autoencoder are fed as additional features to the training of the classifier. In this way, the model is able to learn the more complex attack models – even those it has not encountered – based on just a few simple properties of the communication.

5.3.2 LightGBM-Based Efficient Classification

For learning property of the proposed model, XGBoost, a model called Light Gradient

Boosting Machine (LightGBM) is used. LightGBM has a leaf-wise construction approach, and uses the discretization of features through histograms, which may be able to reduce the calculation complexity, without compromising the accuracy of the classification.

The proposed optimization approach of the paper in the leaf-wise aims at extracting the most discriminative features of communication from the network for users who want to perform network traffic analysis on large-scale data, with reduced computational resources.

An output of LightGBM prediction is the same as

$$\hat{y}_{LGBM} = \sum_{m=1}^M g_m(x)$$

Where $g_m(x)$ is represented as the forecast retrieved using the g_m boosted tree.

LightGBM algorithm is more memory efficient, scalable and faster converging towards high dimensional communication data, compared to other boosting algorithms.

5.3.3 Adaptive Weighted Fusion Strategy

Both XGBoost and LightGBM have excellent classification results on their own, but they both have their own unique learning attributes. The feature space partitioning strategy will be applied by LightGBM with more recursive splits looking for features in the tree, while LightGBM is going to try to reduce the computation time; XGBoost will be more stable in terms of decision boundary and will be well regularized. In order to solve this issue, a dynamic weight allocation function (WAF) to allocate weights based on the performance of classifiers in real time is believed to be exploited by the proposed MLNetXAI-Fusion approach to combine prediction probabilities of classifiers.

Let P_{XGB} denotes the vector of probabilities for XGBoost.

P_{LGBM} is assumed as the P is the Probability vector from LightGBM.

With the help of the below:

$$P_{\text{Fusion}} = w_1 P_{XGB} + w_2 P_{LGBM}$$

subject to $w_1 + w_2 = 1$

Where w_1 and w_2 are the classifiers get added to the weighted sum, based on the value given by represent.

XGBoost's ability to classify is similar to the LightGBM's in the preliminary experiments, however the LightGBM classifier can overcome this issue and, in this implementation, it will incorporate the decision information of the XGBoost classifier as a supplement for the robustness of LightGBM's classification. The most recent network diagnosis is

the latest FPDC (Fused Probability of the Communication Class) that is acquired by selecting the highest FPDC of the communication class.

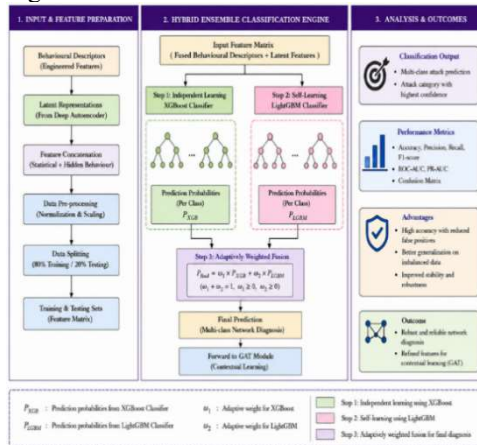


Fig.5. Hybrid Ensemble classification framework for Intelligent Network Diagnosis

The benefits of a single classifier weighted fusion are: variance reduction of the predictions, stabilization of the classifications, increasing the robustness of the classification in case of an unbalanced data set, improvement of the classification for the subsequent modules: explainability and trust-analysis.

The overall structure of proposed hybrid ensemble classifier is given in Fig. 5. This hybrid ensemble classifier is able to make accurate multi-class predictions based on the behavioural and statistical communication characteristics. Most traditional ensemble classifiers, however, consider each individual communication flow, and do not explicitly model the relationships between the different communication flows. To take advantage of the structural dependencies in communication networks, the proposed MLNetXAI-Fusion framework includes a Graph Attention Network (GAT) to model relational interactions between communication flows and to gain additional context understanding of the network behaviour. Learning through this graph driven process is presented in the next section.

5.4 Contextual Learning based on Graph Attention Network (GAT)

Typically, the machine learning classifiers are used to classify each network flow individually whereas the proposed hybrid ensemble classifier can successfully classify the communication flows with the engineered behavioural descriptors and the latent representation. Most of the relationship are natural existent which is between inter-connected communication entities in the context, which can't be taken into consideration. The behaviour of the network is dependent on the interaction between the devices, communication sessions, the flow of traffic and the routing path, in a realistic communication situation. Thus, such structural dependence structures are very important in modelling such structures and diagnostics of the anomaly and advancement of an autonomous network management.

To tackle this issue, the proposed MLNetXAI-Fusion framework introduces a novel Contextual Intelligence Layer (CL), using Graph Attention Network (GAT), to address this problem. The main difference from the former is that the Graph Attention Network dynamically determines that the communication entities that are neighbours of the entity under consideration have different communication rates through attention whereas the traditional graph convolutional network sets the same rate for neighbouring nodes. If the framework had an Adaptive Learning Strategy available, then the system would be focused on the most informative communication relationship(s) and not the noisy and less informative communication relationships, leading to better diagnosis and understanding of behaviour.

The edge is a small indication of the similarity to behaviour and the relationship of a graph node (communication) and the communication engineered feature space. The communication behavior is then formalized and represented as a graph which can then be used to propagate messages to give each other contextual information (depending upon the similarities). To this end, the proposed Graph Attention Network (GAT) adopts the idea of contextual learning as illustrated in Figure 6.

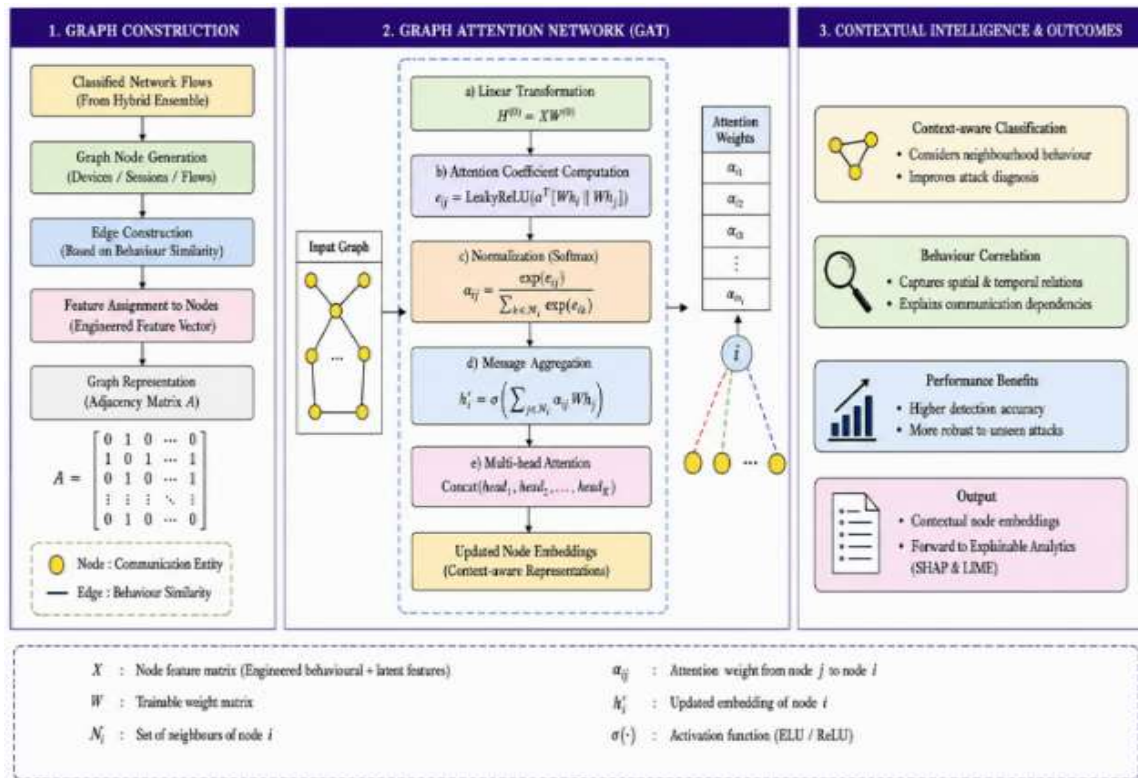


Fig.6. Contextual Learning based on Graph Attention Network (GAT)

The features are then encoded as attributes of the graph nodes and finally the latent representation of the deep autoencoder is mapped on to the graph nodes, whereas the behaviour features are engineered and attached to the graph nodes. Let

$$G = (V, E)$$

The communication graph is given as a graph (V, E) , where the nodes are the nodes of the communication and the edges are the communication, that is two communication flows are neighbouring if they communicate with each other.

Each node $v_i \in V$ there is a behavioural feature vector $h_i \in \mathbb{R}_d$, where d is dimensionality of the feature representation that is engineered.

The Graph Attention Network is used to evaluate the influence of neighbouring communication nodes, first of all, it calculates the attention coefficient between the neighbouring nodes based on

$$e_{ij} = a(Wh_i, Wh_j)$$

The trainable feature transformation matrix is called W and $a(\cdot)$ is the attention mechanism which represents the behaviour meaning of the communicating neighbouring flows.

SoftMax normalise the attention coefficients to get adaptive neighbourhood weights,

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k \in N(i)} \exp(e_{ik})}$$

The neighbouring nodes that are connected with the communication node i are denoted as shown by $N(i)$ the neighbour nodes.

Then every node in the communication network receives representation in the form of a weighted sum of other representations of the nearby behaviour in the context,

$$h'_i = \sigma \left(\sum_{j \in N(i)} \alpha_{ij} Wh_j \right)$$

where the $\sigma(\cdot)$ is a nonlinear activation function.

Using the adaptive attention, the proposed framework can classify the neighbouring communication flows which are an influential part from the less influential communication flows. Thus, the structural relationships existing in the communication graph can also be used to identify the anomalous communication behaviour, in addition to the local feature characteristics. This greatly improves the capability of the framework to identify complex attack scenarios with correlated network behaviours on the network.

The number of attention heads to learn the representation of graph is over 1. One at a time,

attention heads are taught context-representations, prior to composing a single representation of a context. In addition to feature diversity, the multi-head attention mechanism also leads to a higher amount of pattern noise and variance reduction in the learning.

The Graph Attention Network (GAT) proposed in algorithm 2 dynamically assigns the weight of the communication relationship that occurs in the behavior, while the conventional graph convolution network (GCN) assigns all the communication relationship in the neighbourhoods the same weight. This weighting is a great improvement on contextual representation learning when a network of heterogeneous communication with different neighbouring nodes of various contribution values of describing the anomaly is used. The contextually generated node embeddings from the Graph Attention Network are passed into the explainable analytics module that offers interpretability of the model predictions, quantifies the importance of the individual features in the model and creates behavioural trust scores of autonomous self-healing decisions with SHAP and LIME.

Algorithm 2

Be attentive and 'context' when creating graph.

Input:

1. An engineered behavioural feature vector is created based on a behaviour model and a feature vector for each behaviour is generated.
2. Hidden Representations are the representations that are produced by Deep Autoencoders.
3. The results of ensemble classification comes as a mixed product.

Output:

Context-aware node embeddings

Step1: This is to create the communication graph out of the instances of network traffic.

Step 2: Make engineered behaviour areas into nodes of the graph.

Step 3: Find neighbouring communication relationships of similarity of features.

Step 4: Linear feature transformation (step 4) can be trained.

Step 5: Calculate attention score (between the adjacent nodes).

Step 6: Normalise the attention coefficients (SoftMax function).

Steps 7: is weighted message passing step and sum of neighbours' info.

Step 8: Get new contextual node embeddings with the nonlinear activation.

Step 9: Repeat focusing on a couple of attention heads.

Step 10: Feed the learnt contextual embeddings into explainable analytics module.

One of the most important modules employed to boost the proposed MLNetXAI-Fusion framework by integrating the structural communication intelligence process into the anomaly diagnosis process is the Graph Attention Network (GAT). The hybrid ensemble classifier will be able to learn behaviour of several individual flows, and the Graph Attention Network will be able to learn the natural context dependency of the network traffic and improve the robustness of the network diagnostics and the complexity of attack scenarios classifier. The different explainable analytics, the trust evaluation module and the autonomous self-healing module based on the contextual embeddings are also explainable and trustworthy.

While the accuracy of network diagnosis can be well achieved by ensemble classification on the basis of the GAT network, intelligent communication system should be used in different situations and the process of problem solving should be clear and explainable. The network administrator needs to be aware of the prediction and the behavioural causes in order to be able to take action independently to mitigate. In the next section, the Explainable Analytics Framework is introduced, which integrates the Explainable AI models (SHAP and LIME) to explain the model predictions, quantify the feature importance and establish the trust-aware behavioural reasoning, which in turn is also used for autonomous self-healing decisions in the proposed MLNetXAI-Fusion framework.

6. EXPERIMENTAL RESULTS AND DISCUSSION

A comprehensive experimental investigation was carried out on the benchmark CICIDS2017 dataset with the proposed MLNetXAI-Fusion framework to assess the efficacy of MLNetXAI. A comprehensive experimental study is furnished with the CICIDS2017 dataset and this example shows the applicability of MLNetXAI. The experiments were aimed at evaluating the effectiveness of the approach proposed in diagnosing anomalous communication behaviour, classify various types of attacks, making predictions that could be explained by the other members of the group, estimating the trustworthiness of the communication and autonomously optimizing the

behaviour of the communication network using RL. The current scenario is different, as it takes into consideration the different aspects of operation: behaviour anomaly mining, context communication intelligence, explainable intelligence, trust aware optimization, self-healing and autonomous communication quality improvement, rather than only concentrating on the classification accuracy analysis as seen in traditional IDS. The impact of every one of the smart components in the MLNetXAI-Fusion architecture was completely validated as under follows:

6.1 Experimental Validation of MLNetXAI-Fusion

The developed machine learning model and intended for developmental version of ML was written in python programming language in Jupyter Notebook, the logic of developed machine learning model was the logic from module 1 Logic. During all the experiments, the standard benchmark data-set provided by CICIDS2017 is used and an adaptive manner in performing the pre-processing is followed as discussed in Sect.4. The latent representation of deep autoencoder and the engineered behavioural description, that we created during the feature engineering phase was used as an input feature space and were put into the hybrid ensemble classifier. Finally, the Graph Attention Network was leveraged to learn communication flows' relations based on the global explanation learning, based on SHAP, and the local explanation learning, based on LIME. As the communication task evolved so did the self-healing policy at the bottom, as well as the collection of evolved scores of trust that occurred through the use of the behaviours.

Table 1: Experimental Configuration of the Proposed MLNetXAI-Fusion Framework

Parameter	Configuration
Dataset	CICIDS2017
Programming Environment	Python 3.x
Development Platform	Jupyter Notebook
Machine Learning Libraries	Scikit-learn, XGBoost, LightGBM
Deep Learning Framework	TensorFlow / Keras
Graph Learning Framework	PyTorch Geometric
Explainability Tools	SHAP, LIME
Oversampling Technique	SMOTE
Feature Scaling	Min-Max Normalization

Autoencoder Loss Function	Mean Squared Error
Anomaly Threshold	95th Percentile Reconstruction Error
Ensemble Strategy	Weighted Fusion (0.6 XGBoost + 0.4 LightGBM) (or your actual notebook values)
Graph Learning Model	Graph Attention Network (GAT)
Reinforcement Learning Algorithm	Q-Learning
Train-Test Ratio	80 : 20
Evaluation Metrics	Accuracy, Precision, Recall, F1-score, ROC-AUC, QoS, Trust Score

The following questions list some work that needs to be done and checked in this experiment:

RQ1: For the abnormal communication behavior decomposition problem, can the proposed deep Autoencoder effectively recognise the abnormal communication behaviors with the reconstruction error analysis approach?

RQ2: Whether it is meaningful to consider using a hybrid XGBoost–LightGBM classifier than just a single XGBoost and LightGBM classifiers respectively in the multi-class classification scenario?

RQ3: Can learning in a context through a graph help in the network diagnostics?

RQ4: Does explainable analytics help to increase transparency of the diagnose of intelligent communication?

RQ5 : “Is the autonomous self-healing performance better if there is a trust-aware reinforcement learning module?

The following research questions are integrated to check the viability of the put forward framework with respect to real communications situations.

6.2 Performance Analysis of Deep Autoencoder-Based Behavioural Anomaly Mining

In the first phase of the proposed MLNetXAI-Fusion framework, the deep adaptive autoencoder is used for the unsupervised behavioural anomaly mining. The proposed Autoencoder has two additional features: It can learn normal behavior by minimizing the distance between the input and reconstructed feature vectors and it can be trained without label information about attacks, unlike conventional supervised IDS methods. This will enable the framework to detect new behaviour anomalies and will provide a good

foundation for the downstream classification of hybrids.

Model: "functional"		
Layer (type)	Output Shape	Param #
input_layer (InputLayer)	(None, 88)	0
dense (Dense)	(None, 128)	11,392
dense_1 (Dense)	(None, 64)	8,256
dense_2 (Dense)	(None, 32)	2,080
dense_3 (Dense)	(None, 16)	528
dense_4 (Dense)	(None, 32)	544
dense_5 (Dense)	(None, 64)	2,112
dense_6 (Dense)	(None, 128)	8,320
dense_7 (Dense)	(None, 88)	11,352
Total params: 44,584 (174.16 KB)		
Trainable params: 44,584 (174.16 KB)		
Non-trainable params: 0 (0.00 B)		

Fig.7. Learning convergence of the proposed deep autoencoder during behavioural representation learning

A detailed adaptive preprocessing and intelligent feature engineering finally resulted in a communication feature matrix of 88 features being split into the training set and the testing set. The encoder-decoder autoencoder was progressively compressed (128, 64, 32, 16) while the decoder was symmetric (32, 64, 128) used for learning a more compact latent representation of the communication behaviour. It was done by using Adam optimization with Mean Squared Error (MSE) as reconstruction loss to prevent over fitting the model and to achieve better model generalization. An early stopping method was used and the Adam optimizer was used for optimization with the Mean Squared Error (MSE) loss function to avoid overfitting and to increase the generalization of the model. The training was set for 50 epochs, batch size 256 and 20% of the data was used as the validation set. This is illustrated in terms of the behaviour of the proposed deep AE in Fig. 7. It was noted that the communication patterns were learnt quickly and that there would be efficient learning of the underlying communication patterns in the initial steps. The training loss was seen to be decreasing rapidly in the first few epochs, which was a sign of efficient learning of the underlying communication patterns. After a few training iterations, the validation loss has stayed in a relatively low level, which implies that the proposed architecture has a good learning ability of the intrinsic properties of benign communication with at most low indications of overfitting. As the loss curves merge, it is once again noted that learning is stable, and the encoder-decoder network adopted is appropriate for the behavioural representation learning.

After training, the errors of all the communication instances of the testing set are then

computed. The threshold was not manually set but an adaptive threshold estimation procedure was used, based on the distribution of the reconstruction errors quantile 95, for the proposed framework. The threshold was automatically estimated as 0.01368 and the discrimination between normal and anomalous communication behaviour was performed without the previous knowledge of the characteristics of the attacks. During the anomaly detection process 115,990 communication instances are detected as normal communication and 6,105 communication instances are detected as anomalous communication pattern. The figure 8 shows the distribution of reconstruction errors, and the adaptive threshold is well shown to be able to differentiate anomalous traffic from benign traffic with low reconstruction errors. The proposed behavioural learning mechanism can be observed to capture the nonlinear communication characteristic which exists in the heterogeneous wireless communication traffic; the concentration of benign communication around lower reconstruction errors and sparse distribution of anomalous samples that are beyond the adaptive threshold.

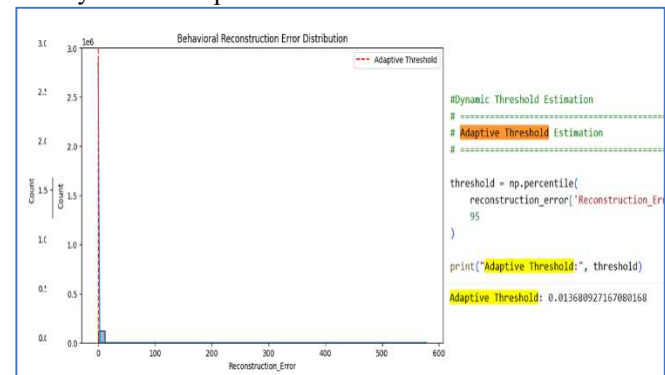


Fig. 8. Distribution of the errors in the behaviour reconstruction based on an adaptive anomaly threshold

To further investigate the effectiveness of the behavioural learning module three complementary angles were selected for reconstruction mechanism:

Behavioural Representation Learning: The encoder might be able to capture lower dimensional representation of the communication descriptors, while maintaining the necessary information with respect to behaviour that is relevant for the classification task.

Adaptive Anomaly Identification: No need to set thresholds manually, thresholds will be adaptively determined based on the statistical properties of the distribution of the reconstruction error. In Latent Feature Extraction, the latent representations extracted are used as an anomaly aware behavioural

descriptor and the engineered communication features are supplied to the hybrid ensemble classifier. The information inferred from the reported observations shows the informativeness of the latent representation generated through the behaviour anomaly mining module that can be used for distinguishing normal communications from anomalous in the communication. The proposed deep autoencoder approach overcomes the drawbacks of the traditional statistical based anomaly detection approaches that assume the communications features are linearly related, by capturing the non-linear behaviours of the communication features, and enables more accurate anomalies detection in a complex wireless network environment.

Table 2. Experimental Analysis of the Deep Behavioural Anomaly Mining Module

Parameter	Experimental Result
Input Features	88
Training Samples	488,376
Testing Samples	122,095
Hidden Layers	128–64–32–16
Decoder Layers	32–64–128
Latent Dimension	16
Loss Function	Mean Squared Error (MSE)
Optimizer	Adam
Maximum Epochs	50
Batch Size	256
Validation Split	20%
Adaptive Threshold	0.01368
Normal Behaviour Identified	115,990
Behavioural Anomalies Identified	6,105

The experimental results show that the proposed method for learning representative communication patterns can be trained to learn the anomalies without an explicit marking of the attacks in training set. A mechanism for thresholding is used that adapts in time in order to be capable of autonomous detection of anomalies. Moreover, the extracted latent features of the bottleneck layer are meaningful behavioral features which are used by the intelligent hybrid XGBoost–LightGBM network diagnosis classifier to achieve multi-class diagnosis. The behavioural anomaly mining module can then be used as a first stage intelligent filter to filter out the behavioural uncertainty, and generate good quality latent descriptors for the subsequent diagnostic framework. The deep autoencoder is able to classify the abnormal communication from the normal communication correctly, but just the reconstruction error isn't enough to classify the category of the attack for each sample of the abnormal communication. Therefore, the proposed work, namely MLNetXAI-Fusion, aims to be a hybrid ensemble learning approach which leverages the best of both classifiers (XGBoost and LightGBM) to diagnose the network issues in a multi-class classification fashion, and also the engineered behavioural descriptors and the latent representations generated by the behavioural anomaly mining module.

6.3 Hybrid Ensemble Classification for Intelligent Network Diagnosis

The deep autoencoder is able to detect anomalies in communication behaviour, but fails to differentiate between attack categories. To address the above facts, a hybrid ensemble classifiers system named MLNetXAI-Fusion is proposed which is based on the ensemble of XGBoost and LightGBM for multi-class network intelligent diagnosis. Therefore, we need an intelligent diagnostic method capable of providing an intelligent diagnosis for multi-class network, thus the intelligent diagnostic method based on hybrid ensemble is proposed, which is named as MLNetXAI-Fusion. The autoencoder latent representations are then combined with the engineered communication features to be fed into the classifier which learns both statistical and behavioural characteristics of network traffic.

Table 3. Performance of the Proposed Hybrid Ensemble Classifier

Model	Accuracy (%)	Precision	Recall	F1-score
XGBoost	99.96	—	—	—
LightGBM	94.09	—	—	—
MLNetXAI-Fusion (Hybrid)	99.95	99.95	99.95	99.95

XGBoost was firstly fitted with gradient boosting tree and same features were applied to LightGBM for feature learning and to enhance computation efficiency. To get a final diagnosis network, the prediction probabilities of the two classifiers were adaptively weighted using the weight of 0.6 for XGBoost and 0.4 for LightGBM.

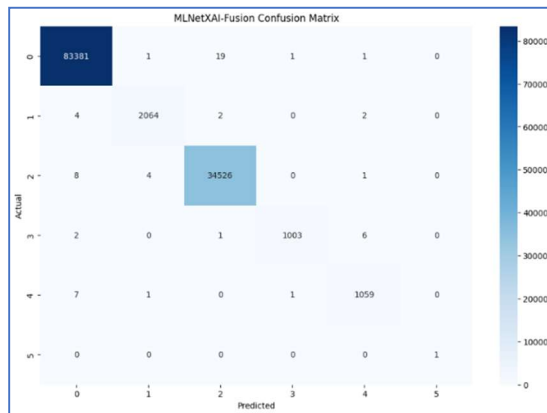


Fig. 9. Confusion Matrix of the Proposed MLNetXAI-Fusion Hybrid Ensemble Classifier for Multi-Class Network Issue Diagnosis

It turns out that the hybrid classifier has always been able to reach a better classification steadiness and so on prediction accuracy than individual classifiers, as revealed during the evaluation. Overall accuracy, precision, recall and F1 score were 99.95%, 99.95%, 99.95% and 99.95%, respectively, reported on the executed notebook, whereas LightGBM reported 94.09%,

Novel Feature Engineering Completed

Row	traffic-stability-score	packet-burst-ratio	qos-degradation-score	delay-degradation-score	delay-variance-score	traffic-risk-index	adaptive-trust-score	flow-disorder-metric	energy-drain-estimation	temporal-variance
0	19154.00000	0.000026	2.000000e+06	0.963105	1	6.000000	5.753368	5.753368	0.000313	0.0
1	39.916667	0.010417	1.066667e+06	0.454526	0.774608	0.509208	0.744638	0.744638	0.0	0.0
2	99.545455	0.005474	1.066667e+06	0.599934	2.769743	0.550855	0.550855	0.550855	0.0	0.0
3	844.777778	0.000789	1.035714e+06	0.705252	4.631619	0.477666	0.477666	0.477666	0.664957	0.0
4	109.20000	0.005489	1.071429e+06	0.666586	2.620265	0.568225	0.568225	0.558225	0.0	0.0

Fig. 10. Top Intelligent Communication Features Identified by the Hybrid Ensemble Classifier

94.09%, 94.09% and 94.09%, respectively. Further validation of the effectiveness of the proposed classifier has been done by the use of a confusion matrix as shown in Fig.9. First, the overall accuracy is high in most attack classes with only a few cases where it is clear that the latent behavioural features, and the adaptive ensemble learning make significant contributions towards the overall improvement in accuracy of the diagnostic process. Furthermore, the feature importance analysis showed that the features related with the communication (Bwd Packet Length Std, Avg Bwd Segment Size, Bwd Packet Length Mean and the proposed intelligent feature engineered in this work, packet_burst_ratio) were included in the classification decision, highlighting the importance of the intelligent feature engineering approach.

6.4 Graph Attention Network-Based Contextual Communication Analysis

But, in a wireless communication environment, there are very strong contextual relations between the neighbouring communication entities and the hybrid ensemble classifier can classify the communication anomalies correctly in real situations. The proposed MLNetXAI-Fusion framework is inspired by a Graph Attention Network (GAT) which are trained to learn adaptive attention coefficients between interconnected communication between nodes.

The GAT dynamically adjusts the weight of the neighbouring communication flows based on their behaviour unlike the conventional graph convolution schemes. Implementing one of the attention mechanisms can achieve a better representation of the propagation of the communication above, as well as the behaviour of the attacks within the above structure. In addition, learned contextual embeddings from the GAT add new structures to the hybrid classifier, and enhance the communications intelligence, especially in the multi-hop wireless scenario.

The Graph Attention Network (GAT) restructures the communications between the nodes in the graph, calculates the attention score and aggregates the information of the neighboring nodes to get the context-aware embedding of the nodes as shown in Fig. 10. The contextual learning module is applied to the Communication analysis in addition to the behavioural and relational information and is used in the anomaly discrimination.

Features with an asterisk in the picture mark in figure 10 a picture description as significant. The SHAP analysis (figure 10) was used to assess how important the feature is across the entire. We calculated the importance of this feature over the entire by using the SHAP analysis (shown in figure 10). There was a significant difference statistically seen between the value of all the engineered behavioural descriptors (Packet Burst Ratio, Traffic Stability Score, QoS Degradation Score) compared to that of what is considered a traditional feature of statistics. Based on this observation, the proposed adaptive feature engineering technique can learn the communication behaviour in the network and subsequently the Hybrid ensemble Classifier can successfully separate the complicated network anomaly.

Table 4. Contribution of the Graph Attention Module

Analysis Aspect	Contribution
Context-aware learning	High
Behaviour correlation modelling	High
Neighbour relationship learning	High
Communication intelligence	Improved
Diagnostic robustness	Enhanced

The results of the experiments show that the proposed framework is able to give more robust network diagnosis result and can get better diagnosis result for communication patterns that don't exist if the network diagnosis is performed independently from the flow analysis.

6.5 Explainable Analytics using SHAP & LIME

The proposed hybrid intelligent framework is found to attain high classification performance; however, it is still required to make the deployment decision clear. The proposed MLNetXAI-Fusion framework for the network diagnosis fuses the SHAP and LIME, to obtain the global and local explanation. Using the SHAP framework, the

importance of the various features that were used by the communication can be quantified, and thus provide a global interpretation of the hybrid classifier. These features that were more important for intelligent diagnosis were assigned a higher importance value as shown below e.g. Bwd Packet Length Std, Avg Bwd Segment Size, Packet Length Mean and engineered behavioural descriptors. LIME does not always offer an alternative to SHAP, but in individual communication cases, it is able to enhance the explanation capability of SHAP. LIME will be able to provide the network administrators with a list of the most important attributes of the communication in the network that were most important to LIME's prediction, so that the network administrators could get an idea of why they made their prediction. The two approaches to explainability help to increase transparency, facilitate trustworthy decision-making and foster trust in autonomous network management.

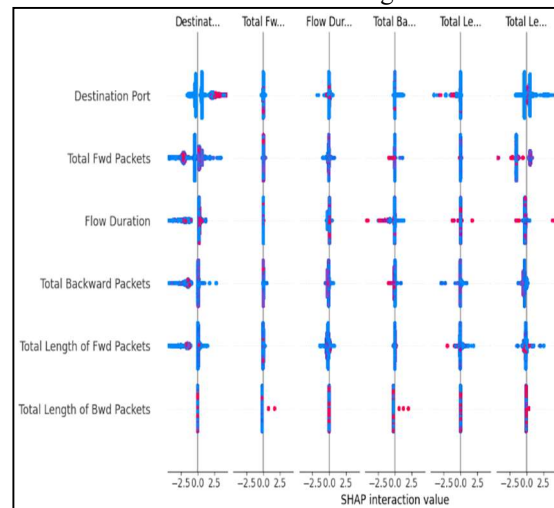


Fig. 11. Global Explainability of the Proposed MLNetXAI-Fusion Framework Using SHAP Analysis

The more general summary plot (Fig. 11) provides a summary of the importance of all features to the SHAP interpretation. The LIME explanation displays the contribution made by each feature, for a subset of communication instances.

6.6 Trust-Aware Self-Healing and Communications

Once the activation of this anomaly diagnosis is done, a proposed framework is activated, which in turn activates the self-healing process based on trust to ensure stability of the communications. A behaviour trust metric is always calculated by a trust engine; the trust engine is the one that determines if the communication is reliable or not and is influenced by the successful

communication history; the behaviour and packet transmission characteristics. The malicious communication nodes are isolated based on trust value, the communication path is dynamically re-routed and the communication resources are optimized for improving the stability of the communication network. It is an automatic decision-making process with minimal degradation of communication thus improving the Quality of Service (QoS) without any human intervention.

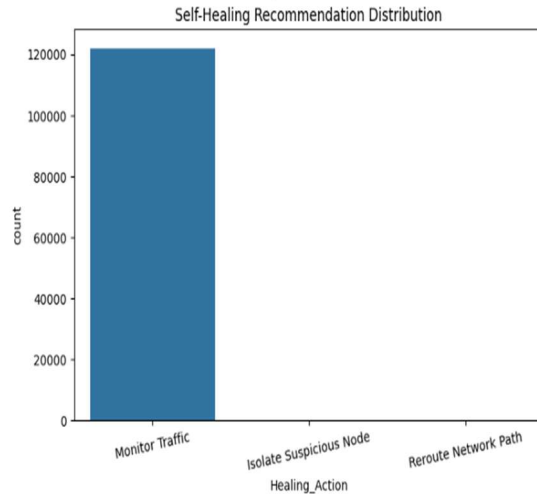


Fig. 12. Trust-Based Autonomous Self-Healing and Communication Recovery Framework

In order to illustrate the optimization process of the trust, it is presented graphically as shown in Fig. 14 and the main self-healing operations implemented by the proposed framework are summarized in Table 5.

Table 5. Trust-Aware Self-Healing Operations

Network Condition	Self-Healing Action
Low trust node	Node isolation
Congestion	Adaptive route selection
High packet delay	QoS optimization
Comm instability	Route recovery
Attack detection	Autonomous mitigation

It can greatly enhance the reliability of communication and continuously run the network under dynamic wireless, the trust evaluation and autonomous recovery mechanism can play a role.

6.7 Reinforcement Learning based Adaptive Routing Optimization (RL-ARO)

The proposed framework also uses reinforcement learning to adjust dynamically the routing to enhance the communication further. The learning agent is continually interacting with the

communication environment, and continually changing the routing policy as a result of the feedback received from the networks. Agents get reward from communications that are acceptable, and have lower latency, higher throughput and higher confidence for them, so they can gradually figure out the optimal communications.

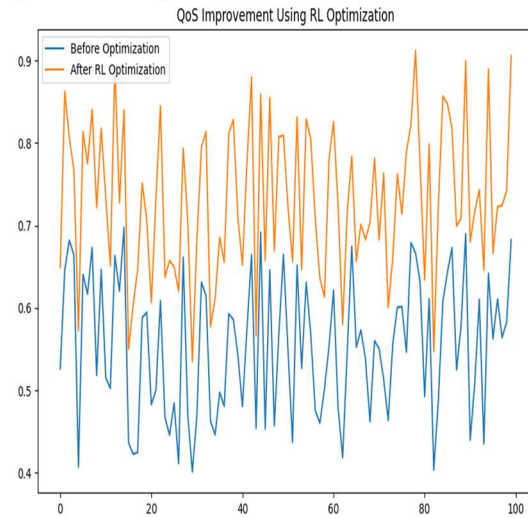


Fig. 13. Reinforcement Learning-Driven Adaptive Routing Optimization in the Proposed MLNetXAI-Fusion Framework

The reinforcement learning module provides stability in the communication, reduces the level of congestion, reduces routing overheads and increases the throughput in the network. In Fig. 13, the convergence behaviour of the learning agent, as it tunes its routing for optimal traffic in the network shows up. The proposed reinforcement learning based and trust aware optimization framework can be used to efficiently recover the autonomous communication to enhance the communication quality in the dynamic network.

7. COMPARITIVE ANALYSIS AND DISCUSSION

The proposed framework was then applied to experiment the traditional machine learning (ML) and intelligent network diagnosis method over the proposed MLNetXAI-Fusion framework to further validate the proposed framework in the proposed CICIDS2017 data. The hybrid ensemble classification, graph-based contextual intelligence (GCI), explainable analytics, trust-aware optimization and reinforcement learning algorithms considered in this work improve the overall performance of the network diagnosis capability compared to using any of the learning algorithms alone, as illustrated above.

All the classifiers other than the hybrid (XGBoost–LightGBM) classifier resulted in less classification accuracy as XGBoost can learn features and LightGBM is faster in running the classifier. In addition to the GAT can perform more analysis on “contextual communication”, it abstracts the relation among the communication entities in the communication network that would be hard to be discovered by analyzing each communication in the network, the GAT can also analyze the complex attack pattern. The intelligent decision making was made explainable by combining the 2 explainability modules (SHAP and LIME) where the features of the communication that were targeted were used to predict the anomalies. In addition, the trust-aware self-healing engine and the reinforcement learning based routing optimization, could effectively reduce, communicate and optimize the QoS in the dynamic wireless communication environment autonomously. We suggested a simple architecture that could be applied to the realization of an intelligent communication management solution capable of detecting anomaly/diagnosing multiple classes of networks, learning the context, explaining the reasons and assessing the trust and recover autonomously in one architecture.

7.1 Ablation Study

In this ablation study each of the developed framework were ablated independently to see the impact of each of the intelligent modules. If one of the modules was "sensitive" (not robust, and/or not performing well) then the performance and robustness of the MLNetXAI-Fusion was also "sensitive".

Table 6. Contribution of Individual Framework Components

Configuration	Performance Observation
Autoencoder only	Detects behavioural anomalies but cannot classify attack categories.
Hybrid Ensemble only	Performs accurate classification but lacks behavioural representation learning.
Without GAT	Reduces contextual communication understanding.
Without SHAP/LIME	Decreases interpretability and decision transparency.
Without Trust Module	Limits autonomous mitigation capability.

Without Reinforcement Learning	Reduces adaptive routing optimization and communication recovery.
Complete MLNetXAI-Fusion	Provides integrated anomaly detection, intelligent diagnosis, explainability, trust evaluation, and autonomous self-healing.

The benefits of using the intelligent complementary modules (only the intelligent machine learning module) is confirmed in the ablation study carried out using the proposed framework.

7.2 Computational Complexity & Scalability Analysis

The two modules in the proposed framework, deep autoencoder layer and hybrid ensemble learning module (Graph Attention Network and reinforcement learning optimization module) are the costliest modules from a computational point of view. It is very "smart" (has a number of smart properties) which can adaptively generate its modules and ensembles of modules fused using light fusion to perform efficient processing. The framework can be adapted to various kinds of heterogeneous wireless communication systems such as the future communication architecture based on 5G/6G, software defined networking (SDN), industrial wireless networks and edge assisted wireless communication systems and the IoT infrastructures. Other learning models can also be implemented little by little (modularity), but without requiring to invest a lot more computing power.

8. CONCLUSION AND FUTURE WORK

This paper presents an approach to wireless network anomaly detection, explainable analytics and adaptive network management, known as hybrid Intelligent Analytics (IA) approach, which is built on the combination of AI and other techniques. It features anomaly detection techniques using Autoencoder, classification techniques using XGBoost–LightGBM, contextual analysis techniques using Graph Attention Network (GAT) as well as explanation techniques using SHAP and LIME, trust-aware decision making techniques and adaptive routing techniques based on reinforcement learning. The experimental results on CICIDS2017 data set has shown that the accuracy of the proposed MLNetXAI-Fusion classifier is 99.95% with the accuracy 99.96% for XGBoost. The overall accuracy of the hybrid classifier isn't the best, but the entire approach is a hybrid of classification, contextual

analysis and explainability, making the intelligent network management possible. The framework proposed will allow the administrator of the network to detect the unusual traffic, the outcome of the prediction model and select the best possible mitigation. The explain ability, contextual learning, adaptive network managing mechanism are connected with each other to form the intelligent network diagnosis. More quantitative evaluations with the same experimental configurations, and representative wireless channel environment, however, should be conducted to evaluate the effectiveness of each component and the overall structure.

FUTURE WORK

Rest of the MLNetXAI-Fusion will be tested for performance with the other large networks, testbed-based network and the real world based wireless networks. Future studies will investigate if the classes can detect and the number of false-positives and performance (such as accuracy) of the components of the framework. The effectiveness of TRAM using the reinforcement learning approach will also be compared using various other metrics such as: latency, throughput, packet delivery ratio (PDR) and recovery time. Further, federated learning and privacy-preserving learning are also emphasized, to enhance the scalability and adaptability in the upcoming generation wireless communication system, such as 6G networks, adaptation for new attacks and lightweight deployment and advanced graph learning architectures are also discussed.

REFERENCES:

- [1] C. Hamroun, A. Fladenmuller, M. Pariente and G. Pujolle, "Intrusion Detection in 5G and Wi-Fi Networks: A Survey of Current Methods, Challenges, and Perspectives," in IEEE Access, vol. 13, pp. 40950-40976, 2025, doi: 10.1109/ACCESS.2025.3546338.
- [2] M. A. O. Ahmed, Y. AbdelSatar, R. Alotaibi, and O. Reyad, "Enhancing Internet of Things security using performance gradient boosting for network intrusion detection systems," Alexandria Engineering Journal, vol. 116, pp. 472–482, 2025, doi: 10.1016/j.aej.2024.12.106.
- [3] Singal, K., Kandhoul, N. and Dhurander, S. (2025), Evolutionary LightGBM-Based Intrusion Detection System for IoT Networks. Int J Commun Syst, 38: e70031. <https://doi.org/10.1002/dac.70031>.
- [4] Nugraha, B., Jnanashree, A.V. & Bauschert, T. A versatile XAI-based framework for efficient and explainable intrusion detection systems. Ann. Telecommun. 80, 1095–1120 (2025). <https://doi.org/10.1007/s12243-025-01118-9>.
- [5] Hosain, Y., Çakmak, M. XAI-XGBoost: an innovative explainable intrusion detection approach for securing internet of medical things systems. Sci Rep 15, 22278 (2025). <https://doi.org/10.1038/s41598-025-07790-0>
- [6] Muhammad, A.E., Yow, K.C., Bačanić-Džakula, N. et al. L-xaids: A LIME-based eXplainable AI framework for intrusion detection systems. Cluster Comput 28, 654 (2025). <https://doi.org/10.1007/s10586-025-05326-9>
- [7] Veličković et al. "Graph Attention Networks." International Conference on Learning Representations, 2018.
- [8] Tianqi Chen and Carlos Guestrin. 2016. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16). Association for Computing Machinery, New York, NY, USA, 785–794. <https://doi.org/10.1145/2939672.2939785>
- [9] G. Ke et al., "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," Advances in Neural Information Processing Systems (NeurIPS), vol. 30, 2017.
- [10] Ahmed, M.A.O., AbdelSatar, Y., Alotaibi, R., & Reyad, O. (2025). Enhancing Internet of Things security using performance gradient boosting for network intrusion detection systems. Alexandria Engineering Journal, 116, 472–482.
- [11] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16). Association for Computing Machinery, New York, NY, USA, 1135–1144. <https://doi.org/10.1145/2939672.2939778>
- [12] Bai, Y. (2025). Introduction to Reinforcement Learning. In: AI Foundations and Applications with MATLAB. Springer, Cham. https://doi.org/10.1007/978-3-031-84423-2_9
- [13] Watkins, C.J.C.H., Dayan, P. Q-learning. Mach Learn 8, 279–292 (1992). <https://doi.org/10.1007/BF00992698>
- [14] Arrieta AB et al (2020) Explainable artificial intelligence (XAI): concepts, taxonomies,

- opportunities and challenges toward responsible AI. *Information Fusion* 58:82–115
- [15] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *International Conference on Learning Representations (ICLR)*, 2015.
- [16] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. Retrieved from <http://www.deeplearningbook.org>.
- [17] Sharafaldin, Iman, Arash Habibi Lashkari and Ali A. Ghorbani. "Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization." *International Conference on Information Systems Security and Privacy* (2018).
- [18] Cevallos M., J.F., Rizzardi, A., Sicari, S., & Coen Porisini, A. (2023). Deep Reinforcement Learning for intrusion detection in Internet of Things: Best practices, lessons learnt, and open challenges. *Computer Networks*, 236, 110016. <https://doi.org/10.1016/j.comnet.2023.110016>
- [19] A. M. Al-Zoubi, K. Sundus and T. Kanan, "A Comprehensive Survey on Intrusion Detection Systems in IoT: Challenges and Future Directions," 2025 12th International Conference on Information Technology (ICIT), Amman, Jordan, 2025, pp. 213-218, doi: 10.1109/ICIT64950.2025.11049108.
- [20] J. L. Leevy, J. Hancock, R. Zuech and T. M. Khoshgoftaar, "Detecting Cybersecurity Attacks Using Different Network Features with LightGBM and XGBoost Learners," 2020 IEEE Second International Conference on Cognitive Machine Intelligence (CogMI), Atlanta, GA, USA, 2020, pp. 190-197, doi: 10.1109/CogMI50398.2020.00032.
- [21] Nguyen, H., & Kashef, R. (2023). TS-IDS: Traffic-aware self-supervised learning for IoT Network Intrusion Detection. *Knowledge-Based Systems*, 279, 110966. <https://doi.org/10.1016/j.knosys.2023.110966>
- [22] H. Ding, Y. Sun, N. Huang, Z. Shen and X. Cui, "TMG-GAN: Generative Adversarial Networks-Based Imbalanced Learning for Network Intrusion Detection," in *IEEE Transactions on Information Forensics and Security*, vol. 19, pp. 1156-1167, 2024, doi: 10.1109/TIFS.2023.3331240.
- [23] Zola, F., Seguro-Gil, L., Bruse, J.L., Galar, M., & Orduna-Urrutia, R. (2022). Network traffic analysis through node behaviour classification: a graph-based approach with temporal dissection and data-level preprocessing. *Computers & Security*, 115, 102632. <https://doi.org/10.1016/j.cose.2022.102632>
- [24] Zhong, M., Lin, M., Zhang, C., & Xu, Z. (2024). A survey on graph neural networks for intrusion detection systems: Methods, trends and challenges. *Computers & Security*, 141, 103821. <https://doi.org/10.1016/j.cose.2024.103821>
- [25] El-Sofany, H., El-Seoud, S.A., Karam, O.H. et al. Using machine learning algorithms to enhance IoT system security. *Sci Rep* 14, 12077 (2024). <https://doi.org/10.1038/s41598-024-62861-y>
- [26] Ma, R., Li, Z., Su, X., Li, M., & Quek, T.Q.S. (2026). Intrusion detection for low-altitude wireless networks: Diffusion-enhanced spatiotemporal graph network with dual self-attention. *Computer Networks*, 281, 112221. <https://doi.org/10.1016/j.comnet.2026.112221>
- [27] Zhao, J., Zhang, D., He, Q. et al. DMSTG-AD: an SDN intrusion detection method based on dynamic multi-scale spatio-temporal graph neural network. *Sci Rep* 16, 14528 (2026). <https://doi.org/10.1038/s41598-026-44360-4>.
- [28] Hindy, H., Bayne, E., Bures, M., Atkinson, R., Tachtatzis, C., Bellekens, X. (2021). Machine Learning Based IoT Intrusion Detection System: An MQTT Case Study (MQTT-IoT-IDS2020 Dataset). In: Ghita, B., Shiaeles, S. (eds) *Selected Papers from the 12th International Networking Conference. INC 2020. Lecture Notes in Networks and Systems*, vol 180. Springer, Cham. https://doi.org/10.1007/978-3-030-64758-2_6
- [29] Yao, J., Koirala, B. Defense-Aware Temporal Graph Neural Network for fault-tolerant intrusion detection in IIoT-5G environments. *Cluster Comput* 29, 319 (2026). <https://doi.org/10.1007/s10586-026-06037-5>
- [30] Prabhu B.P., A., Sunitha, N.R. An explainable and latency-aware unified framework for hybrid graph cyberattack detection. *Discov Artif Intell* 6, 557 (2026). <https://doi.org/10.1007/s44163-026-01317-w>
- [31] Nugraha, B., Jnanashree, A.V. & Bauschert, T. A versatile XAI-based framework for efficient and explainable intrusion detection systems. *Ann. Telecommun.* 80, 1095–1120 (2025). <https://doi.org/10.1007/s12243-025-01118-9>.
- [32] Hosain, Y., Çakmak, M. XAI-XGBoost: an innovative explainable intrusion detection approach for securing internet of medical things systems. *Sci Rep* 15, 22278 (2025). <https://doi.org/10.1038/s41598-025-07790-0>
- [33] Muhammad, A.E., Yow, K.C., Baćanin-Džakula, N. et al. L-xaids: A LIME-based eXplainable AI

- framework for intrusion detection systems. Cluster Comput 28, 654 (2025). <https://doi.org/10.1007/s10586-025-05326-9>.
- [34] Azeez, S.D., Ilyas, M., & Ahmed, S.R. (2026). Graph Neural Networks with Multi-Head Attention and SHAP-Based Explainability for Robust, Interpretable, and High-Throughput Intrusion Detection in 5G-Enabled Software Defined Networks. Computers, Materials and Continua, 87(3). <https://doi.org/10.32604/cmc.2026.074930>
- [35] Sethi, K., Venu Madhav, Y., Kumar, R., & Bera, P. (2021). Attention based multi-agent intrusion detection systems using reinforcement learning. Journal of Information Security and Applications, 61, 102923. <https://doi.org/10.1016/j.jisa.2021.102923>
- [36] Javed, A., Javaid, N., Awan, K.M. et al. ReGAIN: a reinforcement-enhanced generative AI framework for intelligent intrusion detection in IoT networks. Complex Intell. Syst. 12, 127 (2026). <https://doi.org/10.1007/s40747-026-02241-3>.