2025 IMEKO TC-6 International Conference on
Metrology and Digital Transformation - **M4DConf 2025**
Benevento, Italy, September 3-5, 2025

Lightweight Passive Monitoring for Soft Anomaly Classification in Wired Networks on Resource Constrained Microcontrollers

Prabin Dhakal[1], Francesco Picariello[2], Basanta Joshi[3], Nanda Bikram Adhikari[4*]

[1,3,4]*Institute of Engineering, Pulchowk Campus, Lalitpur, Nepal*
[2]*Univeristy of Sannio, Benevento, Italy*
[4*]*adhikari@ioe.edu.np*

Abstract – **This paper presents a method for detecting soft anomalies in Ethernet cables using time domain signal analysis and machine learning. A dataset consisting of oscilloscope captured signals from a CAT5e cable using passive measurement technique under four controlled scenarios was used. To reduce data dimensionality while preserving statistical characteristics, a histogram-based feature extraction process was applied prior to classification. Classification was performed using Decision Tree, Random Forest, and Support Vector Machine (SVM) algorithms under various downsampling rates. Models performances were good at generalizing and predicting the classes using those features. The results demonstrate that histogram features are effective in distinguishing between different anomaly types, even at significantly reduced sampling rates showing potential for real time implementation on low powered microcontroller platforms.**

I. INTRODUCTION

Cyber Physical System (CPS) are an integral part of industry 4.0 and 5.0. CPSs are intelligent systems that closely integrate cyber and physical processes, and exchange data and information in real time [1]. While these systems significantly improve the efficiency, reliability, and functionality of the infrastructure, they also introduce new security challenges that must be addressed [2]. CPSs operate at the intersection of the digital and physical worlds. Thus, vulnerabilities in the cyber domain can have a direct impact on physical operations [3].

Effectively monitoring physical layer of network helps detect issues such as signal degradation, physical damage, and unauthorized access with tapping [2]. The targets of physical layer monitoring for cable communication security are the following: 1) detection; 2) localization; and 3) classification of the anomalies. The type of anomalies can be classified as follows: 1) hard and 2) soft [4]. Hard anomalies are severe and typically cause complete failure of the cable. These anomalies are characterized by physical damage or significant degradation that interrupts the electrical or signal continuity (e.g., open and short circuits) [4]. Soft anomalies are less severe and usually cause degradation in performance rather than complete failure. These anomalies can result in intermittent issues or degraded signal quality (e.g., insulation damage, tapping, and connector issues) [4].

Reflectometry is a technique for diagnosing cable conditions, which exploits the signal characteristics of the reflections of a transmitted signal in the wireline channel [5]. It is beneficial for identifying anomalies, discontinuities, and physical degradation within cables and connectors [6]. The traditional reflectometry techniques are active, i.e., based on injecting a test signal into the cable under test and analyzing the reflected waveforms in different domains [5]. Based on the type of incident signal and the considered domain of analysis, reflectometry can be divided into time-domain reflectometry (TDR), frequency-domain reflectometry (FDR), and time-frequency-domain reflectometry (TFDR) [6], [7].

However, these active methods require interrupting the operability of the communication system. To overcome this limit, passive methods have been proposed by considering link quality indicators, such as bit error rate, state transition, eye pattern, and signal-to-noise ratio (SNR) [6]. These methods can detect partial or full open circuits or short circuits within the cable. Nevertheless, they cannot provide good information about soft anomalies (e.g., insulator degradation) [8]. Kallel et al. [8] proposed a machine learning (ML)-based and cost-effective system for a distributed monitoring solution, which can detect anomalies and their types using a histogram-based technique. However, the cable anomalies were only open/short circuits faults. The research by Balestrieri et al. [5] proposed a new method for detection and diagnosis of three types of soft anomalies using different types of features obtained from samples acquired at several sampling rates. This research is the extension of the work by [5] for lightweight diagnosis mechanism.

The preprocessing using frequency-based measure is always demanding for computation. The research aims to solve this problem by using simpler feature extraction mechanism using time domain analysis that can be further implemented in a low powered micro controller device.

In this paper, Section II covers the theoretical background of impact of soft anomalies in acquired signal from cable. Section III reviews related works on current and past state of art methods on active and passive anomaly detection. Section IV details the proposed methodology, including dataset, feature extraction, and machine learning classifiers. Section V presents results, comparing classifier performance under various scenario and assessing deployment feasibility on microcontrollers. Section VI concludes the research's findings and, Section VII presents current work limitation and suggests directions for future improvement.

II. RELATED THEORY

Different faults have different impacts in signal in time domain. High-impedance tapping introduces minimal load on the transmission cable, causing small signal reflections due to impedance mismatches. Such faults often go undetected by conventional protection systems [9]. Their behavior is described by the telegrapher's equations:

$$\frac{\partial V(x,t)}{\partial x} = -RI(x,t) - L\frac{\partial I(x,t)}{\partial t} \quad (1)$$

$$\frac{\partial I(x,t)}{\partial x} = -GV(x,t) - C\frac{\partial V(x,t)}{\partial t} \quad (2)$$

These equations model voltage and current along the line, accounting for resistance (R), inductance (L), capacitance (C), and conductance (G) [10].

Water ingress increases the local dielectric constant, raising capacitance and reducing the characteristic impedance Z_0:

$$Z_0 = \sqrt{\frac{R + j\omega L}{G + j\omega C}} \quad (3)$$

Air exposure can cause oxidation or insulation degradation, leading to minor parameter shifts. Though subtle, such effects may accumulate and degrade signal quality. Higher-order statistical features and time-varying models are often employed to detect these anomalies [11].

III. RELATED WORKS

The passive measurement is the process of obtaining the measurement statistics of cables conditions without injecting test signals into the channel.

In [2], a passive, noninvasive monitoring system is proposed for Fieldbus networks to detect unauthorized physical access by identifying changes in network impedance caused by intrusion devices (i.e., impedance-matching tapping). Wang et al. [2] achieved a classification accuracy of approximately 99% in detecting device intrusion across three positions along the network.

The research [8] used a histogram to analyze the received signal of fast Ethernet cables using STM32-based microcontroller. Two anomalies, open and short circuit were accurately identified with Classification accuracy of 100% [8].

The research [5] used the passive measurement for diagnosis of the cable anomalies using frequency and time domain based features. The soft anomalies were tapping, water exposed and air exposed, which were introduced on CAT5e cables [5]. They created many features and trained the Decision Tree classifier with different sampling rates ranging from 62.5 MHz to 6.25 GHz. The classification accuracy at 312.5 MHz was found to be 99.05% [5].

Many literatures have worked on different active measurement methods to detect the anomalies. However, the anomalies detection using passive method are limited. In the research [8] only the detection of hard anomalies is carried out using histograms of signals. The research [5] conducted detection and classification of soft anomalies using frequency and time domain features which are much complex and computationally expensive mechanism that can be beneficial for powerful and bench setup but is not suitable for handheld or low powered devices.

IV. METHODOLOGY

A. Dataset

The dataset used in this study is based on CAT5e 4PR cables and includes four classes: normal, air exposed, water exposed, and tapped cables. Anomalies were introduced by altering insulation lengths and introducing tapping branches. Data acquisition was performed using an oscilloscope, capturing 10,000 time-domain records per class over a 5.6 μs window. Each record has 35,000 samples in a single time window. The dataset comprises 9913 records for normal cables and approximately 9600-9700 records for each type of anomaly, as summarized in Table 1. Full details of the dataset and collection methodology are available in [12].

Table 1. Dataset Summary

Anomaly Type	**Nos. of Records**	**Anomaly Lengths**	**Count**
Normal	9913	N/A	1000
Air-Exposed	9632	5 to 50 cm (5 cm steps)	1000
Water-Exposed	9684	5 to 50 cm (5 cm steps)	1000
Tapped	9684	3 to 15 m (3 m steps)	2000

Fig. 1 and Fig. 2 are histograms of acquired signal for a CAT5e cable, at 6.25 GHz and 312.5 MHz respectively. The x-axis represents voltage bins (time-domain features)

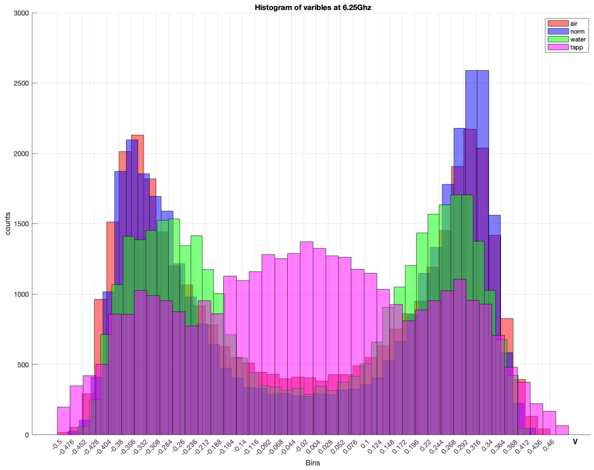

Fig. 1. Histogram Counts of signals at 6.25 GHz

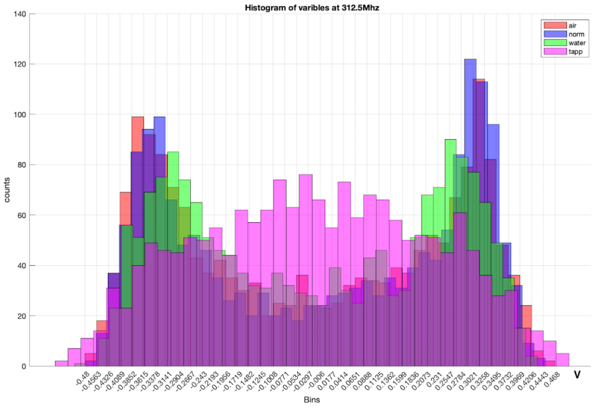

Fig. 2. Histogram Counts of signals at 312.5 MHz

and the y-axis displays the count of occurrences. Four cable conditions types Water exposed, Tapped (Ttap), Air exposed (Tair), and Normal were taken for detection of anomalies in cable. The Water exposed cable curve shows distinct peaks around voltage bins -0.33 V and 0.3 V, indicating significant deviations from the expected -0.38 V and 0.34 V seen in the normal cable signal curve. These shifts might be due to abnormal behavior or unexpected fluctuations in the cable's performance due to water exposure which might be due to change in impedance (Z_0). In contrast, the Tapped (Ttap) cable condition maintains a consistently low and stable count across voltage bins, lacking the pronounced peaks observed in the normal cable curve which might be due to several factors like signal reflection caused by impedance mismatch. This pattern suggests lower sensitivity to signal variations but potential vulnerability due interception. The Air-exposed (Tair) signal displays moderate peaks, with some overlap with both the Water-exposed and Normal cable signals, indicating a potential correlation or shared influence among these conditions. The distribution of the Water-exposed, Air-exposed, Tapped, and Normal cable signal curves across a broader range of voltage bins reflects increased variability which enhances the potential for anomaly detection using machine learning based techniques.

B. *Feature Extraction*

To reduce data dimensionality and retain essential signal characteristics, a decimation process is applied prior to feature extraction. Decimation consists of low-pass filtering followed by downsampling of acuquired signal. Given a discrete-time input signal x[n], the filtered output is computed as:

$$x_f[n] = \sum_{k=-\infty}^{\infty} h[k] \cdot x[n-k] \tag{4}$$

where h[k] is the impulse response of the low-pass filter.

The filtered signal is then downsampled by a factor r:

$$y[m] = x_f[m \cdot r] \tag{5}$$

To avoid aliasing, the filter cutoff frequency is set to

$$(w_c) = \pi/r \tag{6}$$

The decimated signal is subsequently processed using histogram binning, where bin counts serve as features. This approach compresses the dataset while preserving its statistical distribution, enabling efficient and informative feature representation which results in reduced burden to classification process too.

C. *Classification*

The dataset was partitioned using a standard 80/20 train-test split prior to classification. Three machine learning classifiers Random Forest (RF), Decision Tree (DT), and Support Vector Machine (SVC) were evaluated across various downsampling rates, ranging from 31.25 MHz to 6.25 GHz.

For each of the machine learning algorithms, classification performance was assessed using confusion matrices and other performance metrics. Additionally, a training size comparison was conducted to evaluate the adequacy and representativeness of the dataset.

V. RESULTS AND DISCUSSION

The Table 2 shows the performance of three machine learning classifiers, Random Forest, Decision Tree, and Support Vector Classifier (SVC) evaluated across various downsampling factors. These factors represent the rate at which the original data, sampled at 6.25 GHz, is reduced.

Table 2. Performance comparison (validation accuracy) of Random Forest, Decision Tree, and SVC for different downsampling factors

Factor	Rate(MHz)	RF	DT	SVC
1	6250	99.99	99.69	99.99
2	3125	99.99	99.69	99.99
4	1562.5	99.96	99.5	99.99
8	781.25	99.93	99.34	99.97
10	625	99.97	98.32	99.97
20	312.5	99.33	97.37	98.94
200	31.25	79.03	61.10	79.59

Low Downsampling Factors (1, 2, 4, 8): The classifiers demonstrate consistently high performance with accuracies exceeding 99%. This indicates that the essential features of the signal are preserved at these effective sampling rates.

Moderate Downsampling Factors (10, 20): A slight drop in accuracies is observed, especially for the Decision Tree classifier. This suggests potential loss of discriminative features due to smoothing and reduction of details as the sampling rate decreases, affecting simpler models like decision tree more significantly than other two models.

Extreme Downsampling (200): At an effective sampling rate of 31.25 MHz (6.25 GHz / 200), the performance of all classifiers drops significantly. The Decision Tree, a relatively simple model, suffers the most, while the SVC shows comparatively better resilience. This highlights the challenge of feature preservation at such low sampling rates. But still the performance is acceptable for SVC and Random forest classifier.

The drop in classification performance could have been due to the reduction in the number of samples in each record. For instance, at the extreme downsampling rate of 31.25 MHz, each record is reduced to approximately 175 samples, in contrast to the original 35,000 samples captured at 6.25 GHz. This significant reduction in temporal resolution could have impacted classification accuracy due to loss of critical signal features. More detailed analysis is required to quantify this effect. A potential mitigation strategy could involve increasing the time window for signal acquisition to preserve sufficient feature representation even at lower sampling rates.

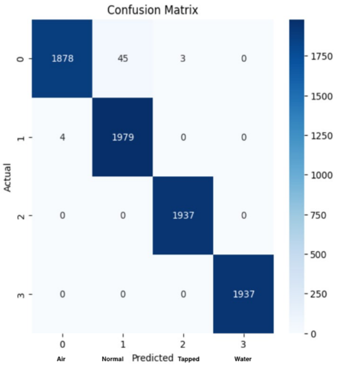

Fig. 3. Confusion Matrix at 312.5 MHz using Random Forest Classifier

Table 3. Classifier overall performance at 312.5 MHz using Random Forest

Class	Precision	Recall	F1-Score	Support
Air(0)	1.00	0.98	0.99	1926
Normal(1)	0.98	1.00	0.99	1983
Tapped(2)	1.00	1.00	1.00	1937
Water(3)	1.00	1.00	1.00	1937

Metric	Score
Accuracy	0.99
Macro Average	0.99
Weighted Average	0.99
Validation Accuracy	0.9933

Model Performance: The confusion matrix (see Fig. 3) reflects the Random Forest classifier's ability to distinguish between four classes: Air exposed, Normal, Tapped, and Water exposed cable as class 0-3 respectively at 312.5 MHz sampling rate. The near diagonal structure of the matrix indicates high true positive rates for all classes. The misclassifications are minimal expect, some Air exposed classes were misclassified as Normal cable, likely due to similarities in signal features as seen in Fig. 1 and 2. Table 3 quantifies the classifier's performance per class using precision, recall, and F1-score for futhuer evaluation of classifier performance at 312.5 Mhz. Classes like Tapped cable and Water exposed cable have perfect scores, suggesting that these anomaly classes produce distinct, consistently identifiable patterns. The slightly lower recall for Air exposed cable may be due to signal patterns that resemble Normal cable behavior under most conditions. Nevertheless, all scores are above 0.98, demonstrating that the model performs balanced and reliable distinction between all the classes. The model globally performs with an overall accuracy of 99.33%, along with high macro and weighted averages. These results confirm that the Random forest classifier generalizes well across all anomaly types.

At a sampling frequency of 312.5 MHz-6.25 GHz, this time-domain based classification achieved a validation accuracy higher than 99% using a Random Forest classifier which is comparable even better than [5] for soft anomaly detection. The support vector based classifier has similar accuracies at different sampling rates comparing to Random Forest. However, the Decision Tree classifier demonstrated worse performance with extremely reduced sampling size. The metrics of classification also validate the feasibility of deploying this lightweight model on low-powered microcontrollers for real-time, passive anomaly detection in wired networks without intrusive probing in communication channel.

Computational complexity: Compared to the approach in [5], which requires a complex O(N log N) computational process,whereas the calculation of histogram of signal level takes only one execution or iteration through each records of the dataset which reflects O(N) complexity for computational resource. It makes the process computationally inexpensive suggesting usability of the feature extraction process even in low powered microcontroller.

VI. CONCLUSIONS & RECOMMENDATIONS

Overall, Random Forest and Support vector machine based classifiers demonstrate robust performance across all downsampling factors, maintaining high accuracy. The results emphasize the trade-off between computational efficiency and model performance when selecting a model for classification of anomalies using histogram based bins as feature. This shows, the potential for using low-powered devices like micro controllers to detect soft anomalies in network cables can be realized by reducing the sampling rate of the signal. By employing time-domain feature extraction using voltage bins histograms as features, the computational requirements for anomaly detection can be significantly minimized. Demonstrating a notable reduction in computational demand during the feature extraction and classification steps.

This suggests that digital conversion devices like lower-bit Analog-to-Digital Converters (ADCs) with reduced sampling rates could effectively detect soft anomalies using time-domain, passive measurement techniques in wired network cables.

VII. LIMITATIONS AND FUTURE WORKS

The use of histograms derived from signal values has been demonstrated as an effective method for classifying cable anomalies, such as air-exposed, high-impedance tapped, and water-exposed conditions. However, the practical implementation of this approach remains limited when relying on the low MHz sampling rates typically found in microcontroller based ADCs. Future work will investigate the impact of extreme downsampling on classification accuracy and explore the use of extended time windows to preserve critical signal characteristics. Real world deployments in environments such as data centers and complex cabled infrastructures present further promising directions for practical validation.

REFERENCES

[1] Zhenhua Yu, Hongxia Gao, Xuya Cong, Naiqi Wu, and Houbing Herbert Song. A survey on cyber–physical systems security. *IEEE Internet of Things Journal*, 10(24):21670–21686, 2023.

[2] Xiangming Wang, Yang Liu, Kexin Jiao, Pengfei Liu, Xiapu Luo, and Ting Liu. Intrusion device detection in fieldbus networks based on channel-state group fingerprint. *IEEE Transactions on Information Forensics and Security*, 2024.

[3] Kang Yan, Xuan Liu, Yidan Lu, and Fanglu Qin. A cyber-physical power system risk assessment model against cyberattacks. *IEEE Systems Journal*, 17(2):2018–2028, 2022.

[4] Qinghai Shi and Olfa Kanoun. Wire fault diagnosis in the frequency domain by impedance spectroscopy. *IEEE Transactions on Instrumentation and Measurement*, 64(8):2179–2187, 2015.

[5] Eulalia Balestrieri, Pasquale Daponte, Luca De Vito, Francesco Picariello, Sergio Rapuano, and Ioan Tudosa. A passive-measurement method for physical security and cable diagnosis. *IEEE Transactions on Instrumentation and Measurement*, 74:1–12, 2025.

[6] Xing-Yu Zou, Hai-Bao Mu, Hao-Tian Zhang, Lan-Qing Qu, Yi-Fan He, and Guan-Jun Zhang. An efficient cross-terms suppression method in time–frequency domain reflectometry for cable defect localization. *IEEE Transactions on Instrumentation and Measurement*, 71:1–10, 2022.

[7] Pasquale Daponte, Gianluca Mazzilli, Enrico Picariello, Francesco Picariello, and Ioan Tudosa. Online diagnosis of automotive wireline channels: the role of measurements and instrumentation. In *2022 IEEE International Workshop on Metrology for Automotive (MetroAutomotive)*, pages 150–154. IEEE, 2022.

[8] Ahmed Yahia Kallel, Dhia Haddad, Thomas Keutel, and Olfa Kanoun. Real-time monitoring of cables based on network interface controllers for predictive maintenance. *IEEE Transactions on Instrumentation and Measurement*, 71:1–8, 2022.

[9] W. C. Santos, F. V. Lopes, N. S. D. Brito, and B. A. Souza. High-impedance fault identification on distribution networks. *IEEE Transactions on Power Delivery*, 32(1):23–32, 2017.

[10] Clayton R Paul. *Analysis of multiconductor transmission lines*. John Wiley & Sons, 2007.

[11] Yatai Ji, Paolo Giangrande, Weiduo Zhao, Vincenzo Madonna, He Zhang, Jing Li, and Michael Galea. Investigation on combined effect of humidity–temperature on partial discharge through dielectric performance evaluation. *IET Science, Measurement & Technology*, 17(1):37–46, 2023.

[12] Francesco Picariello. Two-wire communication channel with anomalies, 2024.