Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 19 No. 13 (2025) | OPEN ACCESS

https://doi.org/10.3991/ijim.v19i13.54539

PAPER

Advanced Anomaly Detection in Mobile Networks: A Hybrid Approach Based on Statistical and Machine Learning Techniques

Meriem Nabil(✉), Meriem Hnida, Abdelhay Haqiq, Imane Hilal

ITQAN Team, LyRica Lab, Information Sciences School, Rabat, Morocco

meriem.nabil@esi.ac.ma

ABSTRACT

Network traffic analysis (NTA) is a technique used by network administrators to monitor network activity, ensure availability, and detect unusual patterns to identify potential anomalies. However, traditional traffic monitoring systems often struggle to detect these anomalies accurately because they rely on rigid models and a limited pool of data. Additionally, anomaly detection is particularly challenging, as anomalies exhibit patterns that differ from most network activities, making their identification based on prior knowledge difficult. This underscores the necessity for an automated and unsupervised approach capable of detecting various types of anomalies despite these limitations. In this paper, we propose an unsupervised framework that combines three statistical machine learning (ML) methods for mobile network failure detection, such as the lower control limit, the cumulative sum algorithm (CUSUM), and the robust stat detector model. Compared to previous studies, models often struggle with large datasets and generate high rates of false positives. However, our approach has proven to be better suited to the demands of traffic monitoring in large telecommunications infrastructures. It offers improved data handling and significantly reduces the rate of false positives while achieving an impressive 98% detection rate for anomalies on telecommunications sites.

KEYWORDS

telecommunication, anomaly detection, change point detection, cumulative sum algorithm (CUSUM), robust stat detector, lower limit control, machine learning (ML), security, cellular wireless networks, predictive maintenance

1 INTRODUCTION

The mobile telecommunications industry began in the early 1970s and has since experienced rapid and unprecedented growth. Starting with the analog "1G" era and moving into the digital "2G" era, the rising demand for mobile technology paved

Nabil, M., Hnida, M., Haqiq, A., Hilal, I. (2025). Advanced Anomaly Detection in Mobile Networks: A Hybrid Approach Based on Statistical and Machine Learning Techniques. *International Journal of Interactive Mobile Technologies (iJIM)*, 19(13), pp. 162–182. https://doi.org/10.3991/ijim.v19i13.54539

Article submitted 2025-01-20. Revision uploaded 2025-04-21. Final acceptance 2025-05-08.

the way for the development of subsequent generations. These advancements have made it possible to access the internet at faster speeds, with technologies such as "3G," "4G," and the current "5G" continuing to push the boundaries of connectivity and performance [1]. With the rapid growth of mobile networks, the number of mobile devices has reached record levels, making security a pressing issue. Users of these devices face increasingly sophisticated threats, including ransomware, spyware, malicious applications, and financial malware.

This technological evolution, while driving innovation and connectivity, has simultaneously provided cybercriminals with unprecedented opportunities to exploit vulnerabilities and target mobile devices on a massive scale. To address this challenge, previous studies have demonstrated that network traffic analysis (NTA) can be an effective method for detecting threats and anomalies [2]. In this perspective, network traffic degradation can be an early indicator of a cybersecurity incident. Many research papers [3], [4] pointed out the relationship between cybersecurity threats and network degradation. For example, attacks such as distributed denial-of-service (DDoS) attempts can overload the network, resulting in network traffic degradation and negatively impacting the quality of service for legitimate users. This highlights the importance of detecting anomalies as early indicators of potential attacks. In this context, the study [5] developed an anomaly-based intrusion detection system by combining two deep learning approaches, convolutional neural networks (CNN) and LSTM. Therefore, proactive detection of network traffic anomalies can help identify and mitigate the effects of cybersecurity attacks, contributing to maintaining the integrity and availability of the network [6].

Indeed, traffic monitoring is essential for detecting abnormal behavior in NTA. This process includes the continuous observation, capture, and reporting of traffic patterns to better understand the nature of the data flow. It helps identify potential security threats, uncovers the sources of bandwidth congestion, and provides insights into overall network performance.

Many tools exist to enable operators to continuously monitor traffic. Machine learning (ML) has become a highly effective approach to traffic detection, thanks to its ability to process large amounts of data, identify patterns, and detect anomalies quickly and accurately. In line with this, the study [7] advocates for integrating ML techniques to enhance the responsiveness of security systems, enabling them to anticipate and prevent anomalies more intelligently and adaptively. However, previous studies [8], [9], [10] in this domain have encountered significant challenges, including a strong dependence on labeled data and the imbalanced distribution of normal and malicious traffic data in real-world scenarios. These challenges not only complicate the development of effective detection methods but also make it difficult to fairly compare the performance of different approaches, because the results of the evaluation may vary depending on the datasets used [11].

Our study stands out by proposing a hybrid and unsupervised approach that combines three statistical detection techniques: the lower control limit, the cumulative sum algorithm (CUSUM), and the robust stat detector. Unlike ML or neural network-based methods, which require large labeled datasets and complex training, statistical methods, such as CUSUM, allow for quick detection of changes in univariate time series, making them ideal for real-time network monitoring. These methods are also highly adaptable, able to handle seasonal variations and traffic spikes, which is crucial in the telecommunications context, where traffic drops are often indicators of anomalies. In contrast to existing methods, our approach also stands out for its ability to generalize and detect a wide range of anomalies related to mobile network degradation, which can be caused by factors such as traffic

congestion, physical barriers, defective equipment, or other factors, while achieving 98% accuracy and reducing false alarms. This flexibility allows it to adapt to the demands of the telecommunications domain and the constraints of large-scale infrastructures, thus offering an effective solution for mobile traffic monitoring.

This paper is divided into seven sections. Section 2 provides an overview of the general context and challenges of network traffic monitoring. Section 3 discusses related work on anomaly detection to identify the existing challenges and issues in the field. By reviewing previous research, this section highlights the approaches, methods, and techniques that have been proposed for detecting anomalies while also emphasizing the limitations and gaps that remain. In Section 4, we outline the applied methodology in this study, which led to the achievement of our results. Section 5 focuses on the practical implementation, detailing the execution steps of the network fault detection models. We then move on to the critical evaluation phase (Section 6), assessing the performance of our network fault detection model to measure its performance and reliability. Finally, we conclude our work and present some future perspectives.

2 CONTEXT AND PROBLEM STATEMENT

In the telecommunications industry, anomaly detection is crucial for enhancing the quality of services. The telecom operators have dedicated multiple departments to monitor their telecom sites. For instance, the network operational center (NOC) is responsible for the network's first-level maintenance and continuous supervision. For instance, once a problem is detected, the NOC takes care of it and contacts the field department, which sends technicians to investigate the problem. If the problem persists, the traffic message channel (TMC) Radio service, which is responsible for second-level maintenance, is contacted. If the problem remains unresolved, the issue is escalated to the supplier for third-level maintenance. While this maintenance process is essential to improving the quality of service, the challenge lies in detecting all anomalies in real-time. Consequently, the NOC plays a critical role in incident detection and is positioned at the top of the operational hierarchy. However, some telecommunications sites may be degraded without the NOC being notified due to various constraints such as physical barriers or defective equipment. In some cases, these issues are not caused by equipment failures but rather by abnormal behavior that impacts network quality, further complicating the situation and highlighting the need for an automated method capable of detecting any type of anomaly despite such constraints.

In this context, our main objective is to develop a robust approach that detects any form of degradation in mobile networks based on mobile traffic by analyzing the average bandwidth consumption of all clients per area. In other words, we aim to identify any anomaly that may signal a potential security issue. Moreover, the fact that detecting such anomalies could reveal early signs of an attack [6] highlights the significance of our work and its potential impact on advancing research in the field of security, as well as contributing to the development of robust systems capable of detecting a wide range of threats.

To achieve this, statistical methods are particularly relevant for detecting mobile traffic anomalies due to their efficiency and quick execution. By offering a balance of simplicity, accuracy, and speed, statistical approaches are well-suited for large-scale telecommunications infrastructures. They also offer greater adaptability when compared to more complex techniques.

In the next section, we provide an overview of related works, examining existing methods and approaches in anomaly detection and their contributions to improving security systems. This analysis will highlight the advancements made in the field, as well as the challenges and limitations that remain to be addressed, allowing us to assess whether statistical approaches truly prove efficient for our context.

3 RELATED WORKS

In this section, we present and discuss several studies that address the problem of anomaly detection. The focus will be on classification-based techniques, distance-based techniques, neural network-based techniques, and statistics-based techniques for detecting anomalies.

3.1 Techniques for detecting anomalies based on classification

This study [12] demonstrates the effectiveness of the isolation forest model (IF) in network anomaly detection, emphasizing its accuracy and low false positive rates. However, it highlights its limitations with large datasets and proposes a parallel algorithm that leverages spark and isolation drill for big data environments. In contrast, Ma, Sun, and Cui [13] discuss the support vector machine classifier (SVM-C) model, which aims to distinguish attack data from normal behavior. Despite its potential, the model's effectiveness is reduced in scenarios involving unlabeled datasets, a common issue in anomaly detection.

In a focused study on zero-day attacks [14], the researchers used CNN and regularization techniques to identify malicious traffic, achieving high accuracy and low false positives with the labeled Bot-IoT dataset. However, their method's reliance on labeled data limits its applicability to scenarios such as mobile traffic anomaly detection without labeled data. Meanwhile, conventional ML methods [15] are challenged by their inability to identify new types of data. Deep learning approaches (DLA) [16], [17], [18] excel due to their superior data learning capabilities. However, deep learning's higher computational demands and reliance on labeled data [19], [20] still pose significant challenges.

Finally, the study [2] mentioned that classification-based methods require the creation of a classification model, which can limit the accuracy of outlier detection. Consequently, distance-based methods were developed with the specific aim of identifying outliers through distance calculations between an object and other objects within a dataset.

3.2 Techniques for detecting anomalies based on distance

The study [21] evaluated five clustering algorithms–k-means, enhanced k-means, k-medoids, expectation-maximization clustering (EM), and distance-based outlier detection–to identify network anomalies signaling potential attacks. These systems detect deviations from standard behaviors, triggering alerts for unusual activities. The study showed that while methods such as K-nearest neighbor (KNN) and Naive Bayes (NB) underperform in scenarios with numerous anomalies, distance-based outlier detection excelled with an accuracy of 80.15%. Nevertheless, a common challenge is the high rate of false positives, which needs to be reduced. Additionally,

ARIAS Luis Antonio Souto noted in this study [22] that high-dimensional spaces reduce the efficiency of distance-based methods due to metric reliance and difficulty in feature interpretation, highlighting a critical area for improvement [23].

Finally, the paper in [24] examines distance-based anomaly detection techniques, with a particular emphasis on memory bank-based approaches and k-distance metrics. It introduces the back to the metric (BTM) method, which utilizes distinct distance metrics for both anomaly detection and segmentation, offering improved performance over the BTF method. However, the BTF method demonstrates weak detection capabilities, despite performing well in segmentation. Additionally, the anomaly score metric requires further refinement to achieve more accurate results.

3.3 Techniques for detecting anomalies based on neural networks

In network security, this paper [25] introduced the denoising auto encoder with generative adversarial network (DAE-GAN), a semi-supervised model designed to extract features from abnormal traffic data. While it proves efficient, the model is heavily reliant on labeled data and may struggle in scenarios with low anomaly levels. Another study [26] tackled the limitations of OC-SVM in handling large datasets and the reliance on pre-labeled data by integrating it with an LSTM auto encoder, improving both anomaly detection accuracy and scalability. Additionally, this paper [27] took it further by introducing a hierarchical deep auto encoder system for phased anomaly detection and classification, reflecting the increasing complexity of network security. These contributions highlight the ongoing efforts to refine anomaly detection with more adaptive, precise, and efficient solutions.

In other research [28], a new algorithm combining CNN, recurrent neural networks (RNN), auto encoders, and generative adversarial networks (GAN) has been proposed to improve anomaly detection in network traffic. This method achieves a 95% accuracy rate in identifying various network traffic anomalies. However, despite its performance, this accuracy is still insufficient for detecting all anomalies, highlighting the need for a more robust tool capable of identifying a wider range of issues.

3.4 Techniques for detecting anomalies based on combining deep learning and statically methods

Reference [29] highlights that while ARIMA models are effective for forecasting, they face challenges in detecting anomalies in time series. It is also noted that hybrid models combining ARIMA and ML have shown promising potential in improving anomaly detection efficiency. In this regard, the study [30] demonstrates the performance of combining LSTM and SARIMA models for predicting and detecting anomalies, overcoming the limitation of fixed thresholds in handling periodic issues. However, this method also identifies traffic spikes as anomalies. In the context of our approach to mobile traffic degradation detection, this identification of spikes as anomalies presents an issue. Indeed, these traffic spikes do not necessarily indicate an anomaly or failure but rather a normal and beneficial activity for the business, marking periods of high demand or operational success. Therefore, classifying them as anomalies could distort the results and lead to errors in the detection process.

3.5 Techniques for detecting anomalies based on statistics

In this study [31], four statistical methods were evaluated for detecting abrupt changes in network traffic. The results revealed that while the K-S algorithm struggled, the X^2 method was effective at identifying traffic drops, mutual information (MI) excelled at detecting traffic spikes, and CUSUM proved to be reliable for both peaks and drops. Additionally, this paper [32] enhances the CUSUM algorithm by integrating wavelet filtering to reduce sensitivity to seasonality, resulting in the WAVE-CUSUM method, which achieves a lower false alarm rate. Meanwhile, Mohammad Braei and associates compared statistical, deep learning (RNN), and classical ML methods (DBSCAN, LOF, etc.) for univariate time series anomaly detection, concluding that statistical approaches are most efficient and effective for identifying anomalies, despite the potential of deep learning approaches in specific contextual anomaly scenarios [33]. Each study contributes to refining anomaly detection strategies, emphasizing the importance of tailored solutions for specific network conditions.

3.6 Comparative Analysis of different techniques in detecting network anomalies

The Table 1 provides a detailed comparison of various strategies for detecting network anomalies, offering a well-organized summary of the models used in this field. It outlines the specific approach of each model, shedding light on the types of anomalies they are designed to detect. Additionally, the table highlights the strengths and weaknesses of each model to analyze their performance, scalability, and practical limitations when applied to network monitoring and security tasks.

Table 1. Benchmarking different approaches for anomaly detection

	Ref.	Models	Advantages	Disadvantages
Based on classification	[12], [13]	Isolation Forest	– High accuracy; – Minimizes false positives.	– Not suitable for large datasets.
	[14]–[20], [2]	OC-SVM, CNN, DLA	– Can characterize a complex border. – DLA models can identify more anomalies than baseline models.	– Training requires a Knowledge Base containing more than normal traffic. – Its low capacity to handle massive datasets. – Very high execution time.
Based on distance	[21]–[24]	K-means	– High accuracy; – Easy to implement.	– Training requires a Knowledge Base containing normal traffic. – The false positive rate is quite high.
		KNN, NB, BTF	– Easy to understand; – Look for the best k.	– Decreased accuracy. – Increased false alarm rate. – Not effective with massive data, as well as its speed declines.

(Continued)

Table 1. Benchmarking different approaches for anomaly detection *(Continued)*

	Ref.	Models	Advantages	Disadvantages
Based on neural networks	[25]–[28]	LSTM, Autoencoder, DAE-GAN, RNN, CNN	– Not sensitive to massive data.	– Training is only done in normal traffic; – Very high execution time; – Average performance in a univariate time series context by detecting only point and collective anomalies.
Combining statistical & DLA	[29], [30]	LSTM-SARIMA	– This combination handle complex, non-linear patterns and periodic fluctuations in data.	– The classification of traffic spikes as anomalies leads to an increase in false positives and classification errors.
Based on statistics	[31], [32], [33]	CUSUM	– Can detect traffic peaks and drops. – Gives better performance on univariate time series by detecting point and collective anomalies. – Requires less runtime.	– Sensitive to seasonality.

Through a comparison of existing works in detecting anomalies in mobile network traffic (refer to Table 1), we found that classification-based approaches, such as IF and OC-SVM, are highly accurate with complex anomaly characterization but less scalable to large datasets and require extensive training. Distance-based methods, including k-means and K-NN, offer simplicity and accuracy, though they struggle with large-scale data and have higher false-positive rates. Neural network techniques, such as LSTM auto encoders and DAE-GAN, handle large data volumes well but are computationally intensive and typically rely on training with normal traffic models. On the other hand, when combining statistical methods with deep learning techniques, such as LSTM-SARIMA, the model becomes more capable of handling complex, non-linear patterns and periodic fluctuations, making it suitable for more intricate data. However, this approach still faces the challenge of misclassifying traffic spikes as anomalies, which leads to an increase in false positives and classification errors. In contrast, purely statistical methods, such as CUSUM, excel in detecting changes in univariate time series with faster execution, although seasonal variations can affect performance. Each approach offers a unique balance of strengths and limitations, reflecting the complexity and diversity of anomaly detection technologies. To summarize, statistical methods prove to be the most suitable, as they provide high accuracy and efficiency while offering the flexibility to categorize either traffic peaks or drops as anomalies. This flexibility is essential in the telecommunications context, where traffic drops are often seen as indicators of anomalies. Based on these conclusions, our proposed approach will focus on the statistical methods.

4 METHODOLOGY

The methodology section outlines the procedure used in this study to develop an effective model for detecting anomalies related to network degradation in mobile telecommunications. The model specifically focused on identifying traffic variations, particularly decreases, within a large set of unlabeled data from 4G telecom sites. As detailed in the previous section, we adopted a statistical approach due to its ability to provide high accuracy, speed, and flexibility in processing large datasets and detecting anomalies.

Initially, we implemented the CUSUM algorithm, known for its ability to detect significant changes in time series data. We apply this model to traffic data from each telecom site, examining the time series of traffic volumes (particularly long-term evolution "LTE") between January 1, 2022, and April 21, 2022. The primary objective of this step is to detect any significant traffic drops that may signal a degradation in network performance. To better handle seasonality and ensure the robustness of the model in detecting anomalies, we integrated seasonal filters that allow the model to adapt to cyclical trends and reduce false positives caused by regular traffic fluctuations.

In the second step, we aimed to address the limitation of CUSUM, which categorizes sites that have experienced temporary degradation as permanently degraded, even if their traffic has returned to normal. To overcome this, we introduce lower control limits, which are specifically designed to detect only sustained decreases in traffic that exceed normal variation thresholds. This approach ensured a more accurate identification of recent anomalies in the telecommunications domain, distinguishing between temporary fluctuations and actual sustained issues.

To further improve detection accuracy, we integrate a third statistical approach to address another limitation of CUSUM, namely its ability to detect only the first traffic degradation, even if it is old and followed by more recent drops. To account for recent trends, we adopted the robust stat detector model. This model complements CUSUM by enabling the detection of multiple simultaneous changes in traffic data. By applying the robust stat detector to the segments identified by CUSUM and the lower control limits, we obtained a more detailed view of long-term traffic degradation, allowing for precise identification of persistent issues and the detection of the beginning of the most recent traffic drop at degraded sites. This combination of statistical methods strengthens network degradation detection, which is then visualized through graphs.

To evaluate the performance of this approach, we treat it as a supervised learning model. To do so, we labeled the output data of our detection method based on the anomalies detected, assigning a value of one to degraded sites and 0 to sites considered normal. This transformation enables the application of binary classification techniques, making model evaluation and optimization more efficient.

In summary, Figure 1 provides a clear understanding of the process followed to develop this solution while ensuring its reproducibility. In the next section, we will delve into the detailed practical implementation of our solution.

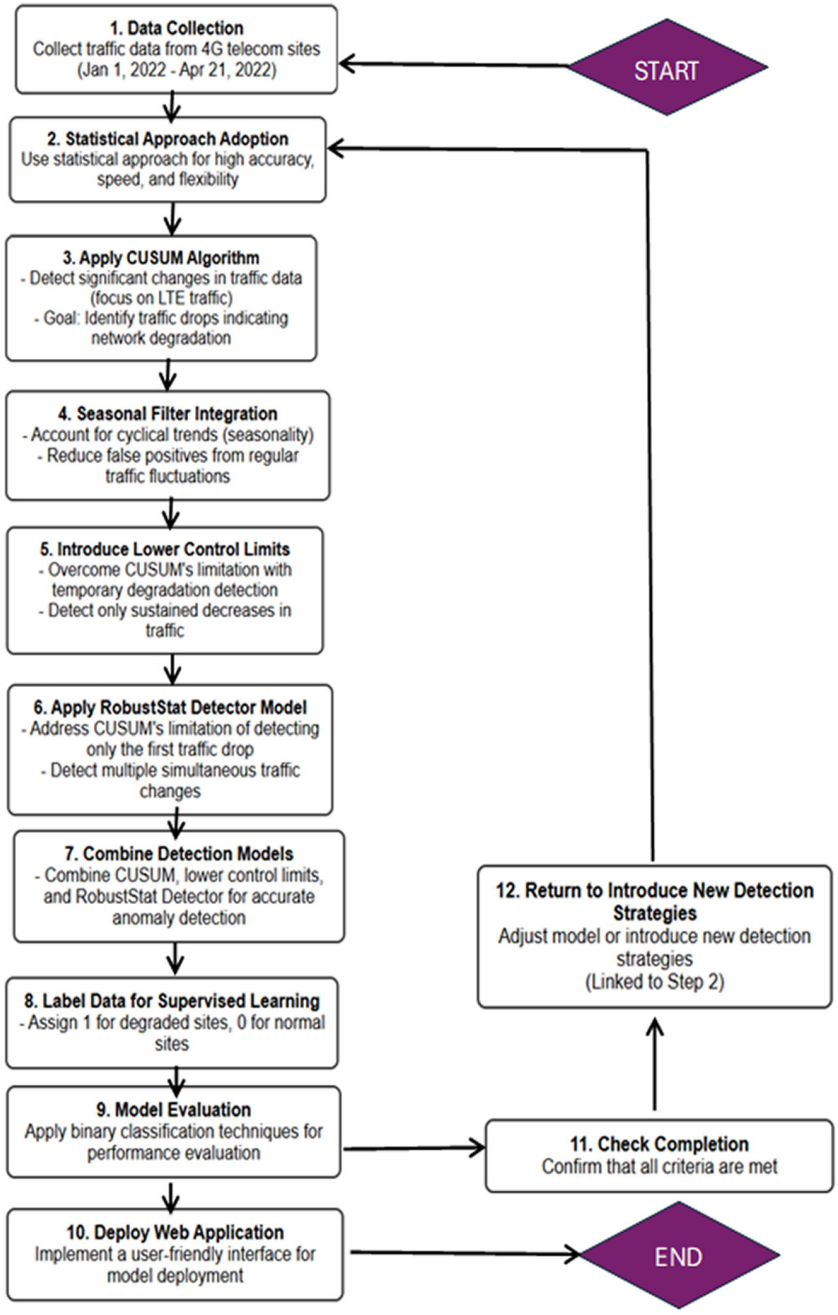

Fig. 1. Flow diagram of our methodology

5 RESULTS

Choosing the statistical approach aligns well with our research, as it effectively detects traffic drops with notable accuracy and speed. To address the CUSUM algorithm's seasonality sensitivity, we will draw inspiration from the work in [32] and integrate seasonality filters to enhance model performance. Additionally, we refine our approach into a labeled model by incorporating classification techniques from [26], utilizing CUSUM's output.

5.1 Dataset features

The dataset contains 4G network traffic data from over 5000 telecom sites recorded every hour from January 1, 2022, to April 21, 2022. The choice of this duration is motivated by the approach adopted in the study [34], which indicates that the CUSUM anomaly detection method may not work perfectly with one-week data. They suggested a month or more because it contains much more information and unexpected events. For the telecom sites classification part, we have added other key performance indicators (KPIs) that may impact the network traffic. We present below ("refer to Table 2") a descriptive table of the elements of our dataset.

Table 2. Presentation of features

Column Name	Description	Type
Day	The day of the transaction	Date
Time	The time of the transaction	Date
Nodebname	The name of the telecom sites	String
Trafic_LTE	4G Traffic Volume (GB)	Double
CSSR	Call Setup Success Rate: % of successful calls.	Double
Indispo	Radio Network Unavailability Rate (%)	Float
Rejet_CSSR_LTE	LTE CSSR Rejection: % of unsuccessful calls.	Float
IPPM	IP Performance Monitoring: the part of the transmission that allows for detecting the voice quality and checking if there is a problem with transmission on telecom sites.	Float

For data preprocessing, the first step was to sort the data by day and time of transaction in order to ensure chronological order, which is essential for the analysis of the time series. We then created a "DAY" column that combines the day and time of the transaction, thus enriching the information available for anomaly detection. One of the main challenges in processing data was the sensitivity of the CUSUM model to seasonality. To address this issue, we have developed a specific function to eliminate the seasonal effect of each site, which has reduced the risk of false positives and false alarms. We used the pmdarima library [35] to estimate the number of seasonal differences needed to make the time series stationary. Seasonal unit tests, such as 'ch' and 'ocsb,' were applied to identify the number of seasonal periods. When two seasonal periods were detected, we differentiated the data over seven periods twice. Through these transformations, in particular the management of seasonality and chronological sorting, data has been optimized for anomaly detection.

5.2 Developing our anomaly detection approach

After the preprocessing of the data of our series, we built our model for the detection of anomalies in mobile networks. In addressing an unsupervised learning challenge, we propose an approach involving three statistical methods: the lower limit, the CUSUM algorithm, and the robust stat detector. In systems with time series, standard measurements such as precision and recall are not suitable for evaluating

anomaly detection algorithms since they don't take into account the specificities of time. To assess the effectiveness of our anomaly detection framework, we utilized the results of our approach to convert the problem into a supervised learning paradigm. By labeling sites as degraded or normal, we can apply algorithms that recognize patterns in the data, allowing better detection and classification of anomalies in mobile networks. This approach allows us to use established measurement values, improving the accuracy and reliability of our detection system. The model's efficiency will be evaluated using a binary classification framework, supplemented by additional KPIs affecting traffic.

Creating the detection model. In our initial detection approach, we employed the algorithm CUSUM introduced by Taylor (2000a) to pinpoint significant changes in the average traffic of telecommunication sites. Using the kats library, we applied the CUSUM model to the NODEBNAME column, treating each telecommunication site individually. The dataset, featuring traffic_LTE and DAY columns, was transformed into a time series after eliminating the effects of seasonality. Focusing on traffic decreases as anomalies, we tailored the implemented approach accordingly. For a clearer understanding of the underlying methodology, the pseudocode of the CUSUM algorithm is provided below, outlining the key steps involved in anomaly detection.

Algorithm 1: CUSUM Algorithm Pseudocode

```
INPUTS:
traffic_data : Dataset of mobile traffic by site
site_list : List of telecom site names to analyze
BEGIN
// Initialization and configuration
FOR EACH site IN site_list DO
// Data filtering and preparation
site_data = Filter traffic_data BY site
// Data preprocessing
processed_data = Remove_Unnecessary_Columns(site_data)
processed_data = Set_Time_Index(processed_data)
// Transform to stationary series
stationary_series = Decompose_Time_Series(processed_data)
stationary_series = Extract_Series_Component(stationary_series)
// Preparation for CUSUM analysis
formatted_data = Rename_Columns_To_Standard_Format(stationary_series)
time_series = Convert_To_Time_Series_Format(formatted_data)
// Apply CUSUM algorithm
detector = Initialize_CUSUM_Detector(time_series)
// Configure to detect only decreases (degradations)
change_points = Execute_Detection(detector, direction="decrease")
// Visualization of results
Generate_Plot(detector, change_points)
Add_Title_And_Labels("Traffic Anomaly Detection for " + site)
Display_Plot()
// Analysis of results
IF change_points NOT EMPTY THEN
Record_Detected_Anomalies(site, change_points)
END IF
END FOR
// Summary of results
Generate_Anomaly_Report(site_list)
END
```

After applying the CUSUM model to identify these anomalies, we visualize degraded sites using Matplotlib (see Figure 2). Figure 2 demonstrates the effectiveness of the CUSUM algorithm in detecting a drop in traffic, which initially leads to the classification of the site as degraded.

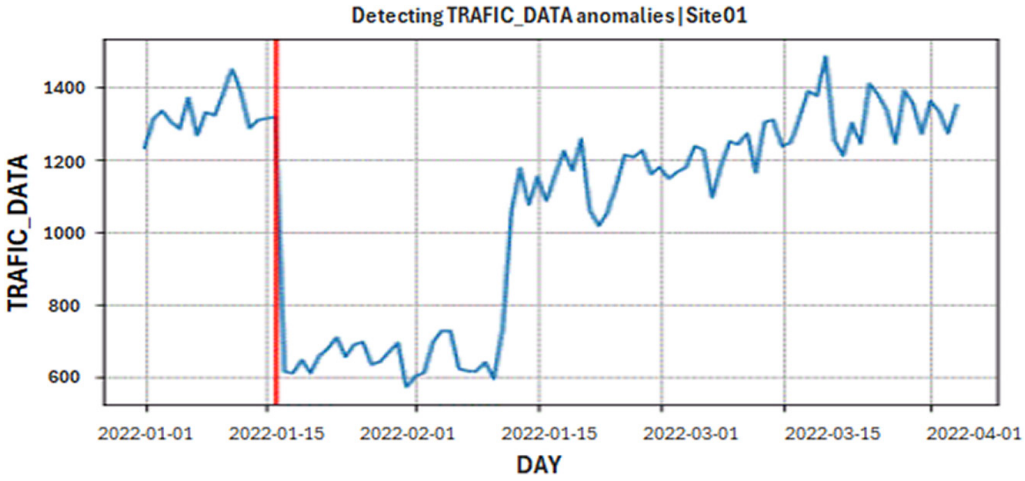

Fig. 2. Visualization of the results of the first detection approach: CUSUM

However, upon further analysis, we observed that the traffic returned to normal levels shortly after the initial drop, indicating that the site was not permanently degraded. This highlights a limitation of the CUSUM approach, as it categorizes temporary fluctuations as persistent degradations.

To address this, we implemented additional methods to more accurately detect real-time anomalies and ensure that temporary fluctuations are not incorrectly classified as sustained issues. According to the study [36], the Shewhart control chart is an effective solution, as it is one of the most powerful tools for monitoring process stability and detecting unusual variations while minimizing false alarms. It is based on the analysis of specific measurements and the definition of control limits, which are statistically determined using an interval based on the mean and standard deviation of the observed data. These limits, known as the lower control limit (LCL) and upper control limit (UCL), help identify significant deviations in the process and detect anomalies, calculated using the following formula:

$$lclimX = \mu X - 3\sigma X$$

$$uclimX = \mu X + 3\sigma X$$

With: data set $\mathbf{X = \{x1, x2, ..., xn\}}$, $\mathbf{\mu X}$ is the mean and $\mathbf{\sigma X}$ is the standard deviation of X.

The main goal of our study is to detect traffic drops, which are often indicative of anomalies caused by various factors such as congestion, physical obstacles, defective equipment, or other issues, while an increase in traffic generally signals positive activity for the telecom operator. Therefore, in this paper, we specifically focus on the lower control limit.

However, instead of using the traditional $\mu X - 3\sigma X$ limit, we chose to apply $\mu X - \sigma X$ to enhance the sensitivity of the detection. In the context of telecommunications, natural traffic variations can be significant, and a $\mathbf{3\sigma}$ threshold might only detect extreme anomalies, overlooking less frequent but still important traffic drops.

By reducing the threshold to **1σ**, we enable the system to more quickly identify significant traffic drops while better tolerating normal fluctuations, thus helping to reduce false alarms.

Here, **μX** represents the mean of the traffic values from the previous four weeks, calculated as the average of traffic values for the same day in each of the last four weeks (traffic value for day *j*-7, *j*-14, *j*-21, and *j*-28). The principle behind this approach is to compare the traffic of the most recent day with the average traffic of the same day over the past four weeks, which helps to account for weekly traffic patterns. **σX** is the standard deviation of the traffic values for control days, providing a measure of the variability in the traffic during these periods, and is expressed as:

$$LCL = AvG(traffic\ value\ j - 7 + traffic\ value\ j - 14 + traffic\ value\ j - 21 + traffic\ value\ j - 28) - Std(control\ day\ traffic\ value)$$

If the traffic on the last day is lower than the calculated average from previous days, and the CUSUM algorithm has already detected a degradation, this suggests that the site's traffic is still significantly below normal levels, indicating a sustained issue rather than a temporary one. In this case, the LCLs act as a check to confirm that the traffic drop is not just a regular variation but a persistent problem fluctuation (see Figure 3). The green line in Figure 3 corresponds to the lower threshold.

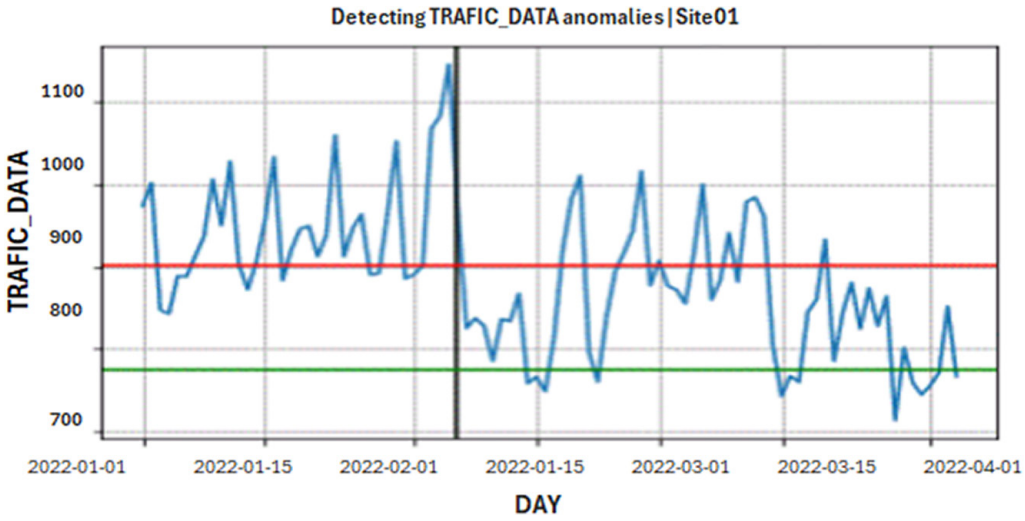

Fig. 3. Visualization of the second detection approach: CUSUM and lower control limit

As shown in Figure 3, site01 has not returned to its normal state; instead, we observe a continuous degradation of traffic over time. This degradation begins at the point detected by CUSUM (black line). This demonstrates that comparing traffic values from previous days helps distinguish between temporary drops and ongoing degradations. This approach ensures that sites flagged as degraded by CUSUM remain classified as such until their traffic returns to normal levels. The integration of LCLs helps reduce false positives, making anomaly detection more reliable by confirming that the site's performance has not recovered and is still experiencing degradation.

Cumulative sum algorithm allows us to detect a single change in network traffic. In Figure 3 above, we observe that degradation begins with the first drop encountered. However, we have been more interested in traffic degradation over the last few days, so we need to incorporate another approach to specify precisely the start

of the last traffic drop. To refine our detection of prolonged degradations, we introduced the robust stat detector model on the segments identified by cumulative sum algorithm.

This model, similar to CUSUM, has the advantage of detecting multiple changes simultaneously. In our approach, we have configured it to identify the latest traffic drop. When applied to the portion of traffic identified as degraded by CUSUM, it allows us to pinpoint the exact moment when the most recent degradation began. If the robust stat detector fails to identify this drop, the last drop detected by CUSUM is considered significant. This integrated approach provides a more detailed assessment of telecom network degradation over time, ensuring more reliable anomaly detection (see Figure 4).

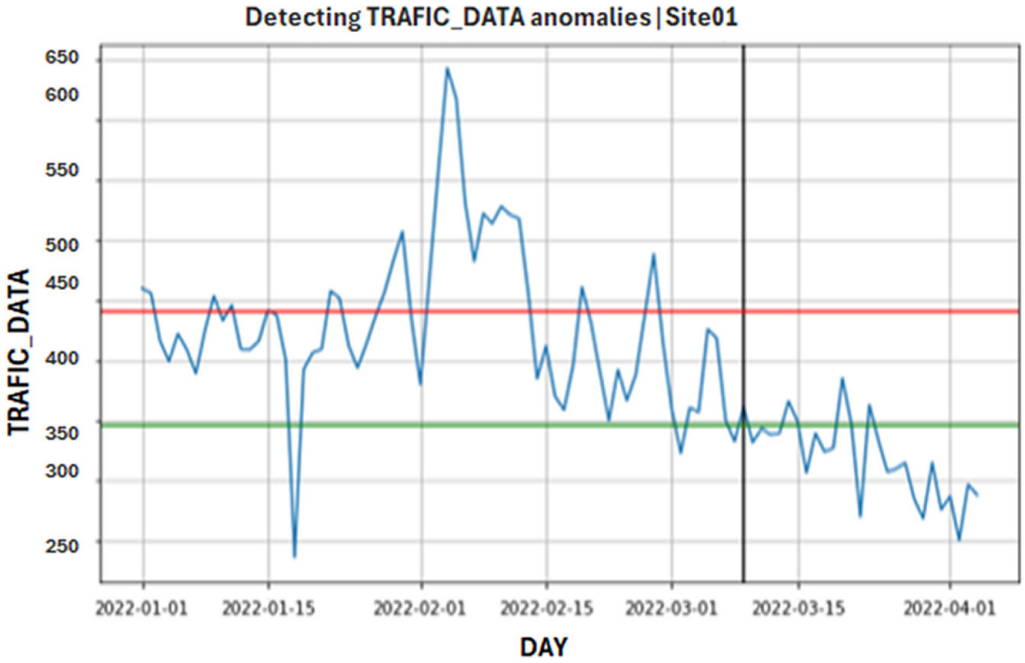

Fig. 4. Visualization of the combined detection approach: CUSUM, lower control limit and robust stat detector

Figure 4 shows the output of our hybrid approach, where we can observe that our method has successfully detected the start date of the most recent traffic degradation, signaling the presence of an ongoing anomaly. The black line indicates the detected degradation, clearly illustrating the traffic variations in relation to the predefined thresholds. The green line represents the lower threshold, and the red line indicates the UCL, calculated as:

$$UCL = AvG(traffic\ value\ j - 7 + traffic\ value\ j - 14 + traffic\ value\ j - 21 + traffic\ value\ j - 28) + Std(control\ day\ traffic\ value)$$

To evaluate our combination of statistical approaches and ensure effective detection of all degraded telecommunication sites, we included an “Anomaly” column in our dataset, with a value of 0 in the absence of an anomaly and 1 otherwise. This transforms our problem into supervised learning, thereby facilitating the measurement of the performance of our approach and assessing its ability to detect all degraded sites through binary classification.

Evaluation steps. In this evaluation step, the goal is to assess the performance of the binary classification model and guide the necessary adjustments to achieve more accurate and useful prediction results. After data preprocessing, including the addition of the "Anomaly" column and handling missing values, we transformed the problem into supervised learning. Features, including additional KPIs, were encoded and used as inputs for the model, with the output representing the presence or absence of anomalies. The data was split into training and test sets (80%/20%), followed by an oversampling step to balance the dataset.

After performing K-fold cross-validation, several models were evaluated, including logistic regression, XGBClassifier, KNN, gaussian, perceptron, linear SVC, SGD, decision tree, and Random Forest. To optimize the performance of these models, we utilized the GridSearchCV library. The results revealed that models such as logistic regression, kKNN, and linear SVC outperformed the others, highlighting the importance of hyper parameter optimization. The application of optimization techniques led to a significant improvement in model performance, particularly for logistic regression. The final model validation was conducted through ROC curve analysis (see Figure 5), with logistic regression standing out as the top-performing model with an accuracy of 98%, confirming its effectiveness in anomaly detection.

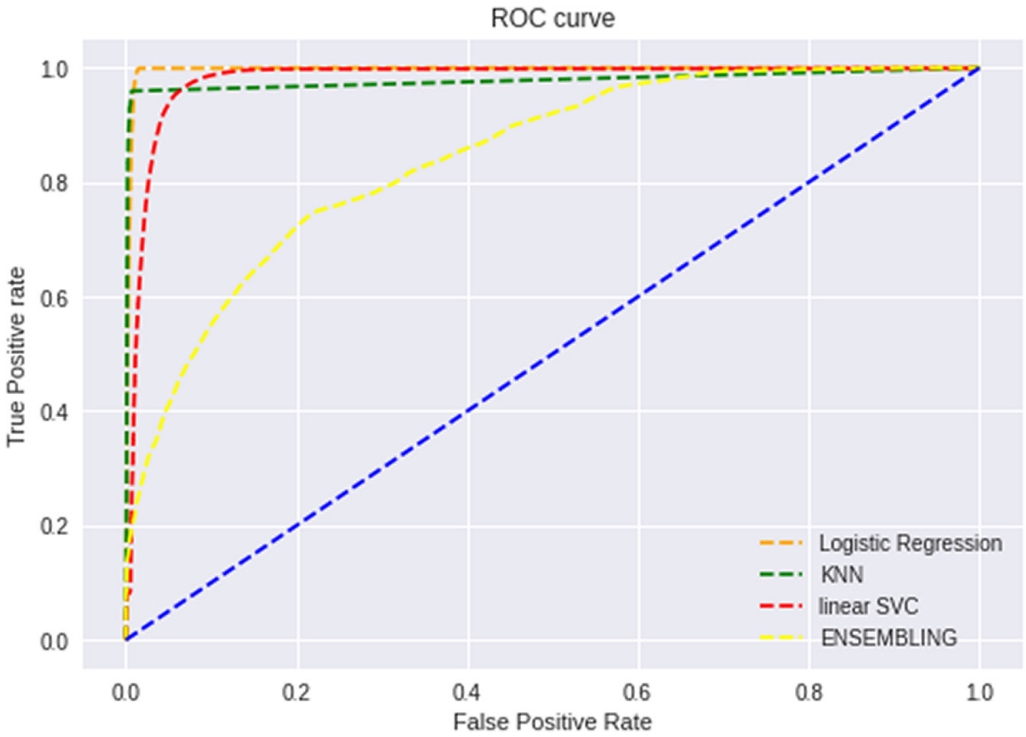

Fig. 5. ROC curve

Figure 5 highlights the receiving operating characteristic (ROC) curve. This curve provides a score by comparing the rate of false positives with the rate of true positives. Plotting the ROC curve allows visual observation of the model's performance. The area under the ROC curve, an additional measure, quantifies the overall quality of the model, with a larger area indicating superior performance. Table 3 displays the performance of each binary classification model based on accuracy, F1 score, and recall indicators.

Table 3. Performance of binary classification models

Approaches	Accuracy	F1 Score	Recall
Logistic Regression	0.98	0.98	0.98
KNN	0.94	0.93	0.87
Linear SVC	0.95	0.94	0.90

Based on the performance indicators (refer to Table 3), we conclude that the logistic regression model is the best-performing model, as it obtained an accuracy of 98% and the recall metric is 98%. These results demonstrate the model's ability to accurately identify anomalies with high precision. In summary, the high AUC and recall values confirm the effectiveness of our detection technique, which successfully achieves an overall accuracy of 98%, validating its reliability for network anomaly detection.

6 EVALUATION

The results show that our model was able to detect 98% of anomalies in mobile traffic, achieving an impressive recall of 98%. This highlights the model's strong capability in identifying the vast majority of real anomalies. As shown in Table 4, existing anomaly detection approaches present trade-offs: IF [12] achieves 87.15% accuracy with a low false positive rate, while CNN [14] improves recall to 99.99% but with an accuracy of 90.75%. Distance-based methods, such as BTF [24], reach 93% accuracy, whereas hybrid models such LSTM-SARIMA [30] achieve 100% recall but at the cost of lower accuracy (81.15%). Unlike these methods, our approach, which integrates CUSUM, the robust stat detector, and the LCL, provides an optimal balance between accuracy and recall, ensuring both effective and reliable anomaly detection.

Table 4. Comparison of our proposed method with other anomaly detection approaches

Approaches based on	Model	Accuracy	Recall
Classification	Isolation Forest [12]	87.15%	– (low false positive)
	CNN [14]	90.75%	99.99%
Distance	BTF [24]	93%	–
Neural networks	CNN-RNN-GAN-Autoencoders [28]	95%	–
Combining statistical and Deep Learning	LSTM-SARIMA [30]	81.15%	100%
Our proposed method	CUSUM-RobustStat Detector -Lower limit control	98%	98%

By integrating the CUSUM algorithm, LCLs, and the robust stat detector, the model efficiently distinguishes between normal fluctuations and actual network degradation. This adaptability is crucial in maintaining a high anomaly detection rate, minimizing false positives, and reducing classification errors. The proposed approach demonstrates robust performance even in scenarios with increased network traffic or high-density telecom environments. The model's dynamic adaptability allows it to adjust efficiently to traffic variations while maintaining its effectiveness in anomaly

detection, even during traffic spikes. Consequently, the method remains reliable and accurate, making it highly suitable for large-scale telecom infrastructures, where traffic spikes and high user density are common challenges.

7 DISCUSSION

To improve the detection of real anomalies while minimizing false positives, we combined three robust methods: the CUSUM algorithm, which identifies significant traffic drops; LCL, which help differentiate temporary drops from sustained degradations; and the robust stat detector, which refines detection by considering multiple simultaneous anomalies, particularly recent traffic drops. Based on our evaluation, our approach has demonstrated an outstanding ability to effectively distinguish normal traffic fluctuations from actual network degradations, achieving 98% anomaly detection and a recall of 98%, which reduces classification errors and maintains high accuracy.

Compared to previous studies, models such as LSTM, OC-SVM, k-means, and K-NN struggle with large datasets and tend to generate high rates of false positives. Our approach has proven better suited to the demands of traffic monitoring in large telecommunications infrastructures. It stands out due to a combination of statistical methods, which have been particularly effective in our context, allowing for the detection of degraded telecommunications sites that are often overlooked by maintenance systems due to specific constraints, such as physical obstacles or equipment malfunctions. It also dynamically adapts to seasonal variations and traffic spikes, ensuring reliable detection even under high traffic conditions. However, while our model is effective, there is still room for improvement, particularly in enhancing precision, further reducing false alerts, optimizing recall, and ensuring stable performance in real-time detection scenarios, which remains challenging in dynamic environments. In this regard, the study [37] highlights that real-time detection remains a significant obstacle, especially due to the limitations associated with managing large and diverse datasets. Therefore, this paper [38] can serve as a foundation and inspiration for future research, particularly in the improvement of our detection approach. Its multi-task learning strategy shows how combining related tasks can enhance accuracy, reduce false positives, and support real-time detection, which are key objectives in our ongoing work. Additionally, the study [39] provides valuable insights for our future research, especially in enhancing anomaly detection systems to achieve more reactive and resilient detection in dynamic environments. This paper underscores the necessity for advanced anomaly detection systems, enhanced authentication protocols, and stronger encryption standards to ensure secure real-time communications.

8 CONCLUSION

In conclusion, this study has enabled the development of a method for detecting anomalies in mobile traffic, with the primary goal of quickly and accurately identifying abnormal values in a time series. Unlike traditional algorithms, which require labeled data and rely on predefined models of normal data, our approach, which combines three statistical methods (CUSUM, LCLs, and robust stat detector), stands out for its ability to analyze unlabeled data. This meets the specific needs of telecom operators, allowing them to detect most network traffic degradations on mobile

networks with 98% accuracy. The results demonstrate a significant impact, providing monitoring services with more responsive tools better suited to the demands of large-scale telecommunications infrastructures.

To make this model more proactive, future work could focus on integrating advanced AI techniques to improve the adaptability and accuracy of real-time anomaly detection. Adapting this anomaly detection approach in real time would enable the system to dynamically adjust to changing network conditions and emerging threats. These advancements could further enhance the efficiency and scalability of the approach, minimizing traffic disruption and providing a robust solution to the growing challenges of mobile network security.

9 REFERENCES

[1] M. Attaran, "The impact of 5G on the evolution of intelligent automation and industry digitization," *J. Ambient Intell. Humaniz. Comput.*, vol. 14, pp. 5977–5993, 2023. https://doi.org/10.1007/s12652-020-02521-x

[2] A. Wahid and A. C. S. Rao, "A distance-based outlier detection using particle swarm optimization technique," in *Information and Communication Technology for Competitive Strategies*, in Lecture Notes in Networks and Systems, S. Fong, S. Akashe, and P. Mahalle, Eds., vol. 40, 2019, pp. 633–643. https://doi.org/10.1007/978-981-13-0586-3_62

[3] A. Privalov, V. Lukicheva, I. Kotenko, and I. Saenko, "Method of early detection of cyber-attacks on telecommunication networks based on traffic analysis by extreme filtering," *Energies*, vol. 12, no. 24, p. 4768, 2019. https://doi.org/10.3390/en12244768

[4] R. Sivaguru, R. Srinath, M. Sathiya Rubha, R. Yasmin Banu, and K. Sathish Kumar, "Network traffic based ransomware detection," *Int. Educ. Res. J.*, vol. 10, no. 3, 2024. https://doi.org/10.21276/IERJ24783683998034

[5] S. Alshattnawi and H. R. Alshboul, "Combined deep learning approaches for intrusion detection systems," *Int. J. Interact. Mob. Technol. (IJIM)*, vol. 18, no. 19, pp. 144–155, 2024. https://doi.org/10.3991/ijim.v18i19.49907

[6] Ł. Wawrowski *et al.*, "Anomaly detection module for network traffic monitoring in public institutions," *Sensors*, vol. 23, no. 6, p. 2974, 2023. https://doi.org/10.3390/s23062974

[7] A. K. Al Hwaitat *et al.*, "Overview of mobile attack detection and prevention techniques using machine learning," *Int. J. Interact. Mob. Technol. (IJIM)*, vol. 18, no. 10, pp. 125–157, 2024. https://doi.org/10.3991/ijim.v18i10.46485

[8] Ankita, S. Rani, and A. K. Bashir, "Analysis of machine learning and deep learning intrusion detection system in internet of things network," in *2022 International Conference on Data Analytics for Business and Industry (ICDABI)*, 2022, pp. 1–9. https://doi.org/10.1109/ICDABI56818.2022.10041259

[9] U. Inayat, M. F. Zia, S. Mahmood, H. M. Khalid, and M. Benbouzid, "Learning-based methods for cyber attacks detection in IoT systems: A survey on methods, analysis, and future prospects," *Electronics*, vol. 11, no. 9, p. 1502, 2022. https://doi.org/10.3390/electronics11091502

[10] T. Ahmad, D. Truscan, J. Vain, and I. Porres, "Early detection of network attacks using deep learning," in *2022 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, 2022, pp. 30–39. https://doi.org/10.1109/ICSTW55395.2022.00020

[11] Z. Chen *et al.*, "Machine learning based mobile malware detection using highly imbalanced network traffic," *Inf. Sci.*, vol. 433–434, pp. 346–364, 2018. https://doi.org/10.1016/j.ins.2017.04.044

[12] X. Tao, Y. Peng, F. Zhao, P. Zhao, and Y. Wang, "A parallel algorithm for network traffic anomaly detection based on Isolation Forest," *Int. J. Distrib. Sens. Netw.*, vol. 14, no. 11, 2018. https://doi.org/10.1177/1550147718814471

[13] Q. Ma, C. Sun, and B. Cui, "A novel model for anomaly detection in network traffic based on support vector machine and clustering," *Secur. Commun. Netw.*, vol. 2021, no. 1, p. 2170788, 2021. https://doi.org/10.1155/2021/2170788

[14] B. I. Hairab, M. Said Elsayed, A. D. Jurcut, and M. A. Azer, "Anomaly detection based on CNN and regularization techniques against zero-day attacks in IoT networks," *IEEE Access*, vol. 10, pp. 98427–98440, 2022. https://doi.org/10.1109/ACCESS.2022.3206367

[15] M. S. E. Sayed, N.-A. Le-Khac, M. A. Azer, and A. D. Jurcut, "A flow-based anomaly detection approach with feature selection method against DDoS attacks in SDNs," *IEEE Trans. Cogn. Commun. Netw.*, vol. 8, no. 4, pp. 1862–1880, 2022. https://doi.org/10.1109/TCCN.2022.3186331

[16] G. G. González, S. M. Tagliafico, A. Fernández, G. G. Sena, J. Acuña, and P. Casas, "One model to find them all deep learning for multivariate time-series anomaly detection in mobile network data," *IEEE Transactions on Network and Service Management*, vol. 21, no. 2, pp. 1601–1616, 2024. https://doi.org/10.1109/TNSM.2023.3340146

[17] K. Hooshmand and D. Hosahalli, "Network anomaly detection using deep learning techniques," *CAAI Trans. Intell. Technol.*, vol. 7, no. 2, pp. 228–243, 2022. https://doi.org/10.1049/cit2.12078

[18] B. Ibrahim Hairab, H. K. Aslan, M. S. Elsayed, A. D. Jurcut, and M. A. Azer, "Anomaly detection of zero-day attacks based on CNN and regularization techniques," *Electronics*, vol. 12, no. 3, p. 573, 2023. https://doi.org/10.3390/electronics12030573

[19] Y. Luo, Y. Xiao, L. Cheng, G. Peng, and D. D. Yao, "Deep learning-based anomaly detection in cyber-physical systems: Progress and opportunities," *arXiv preprint arXiv:2003.13213*, 2021. [Online]. Available: http://arxiv.org/abs/2003.13213

[20] Z. B. Celik, E. Fernandes, E. Pauley, G. Tan, and P. McDaniel, "Program analysis of commodity IoT applications for security and privacy: Challenges and opportunities," *ACM Comput. Surv.*, vol. 52, no. 4, pp. 1–30, 2020. https://doi.org/10.1145/3333501

[21] I. Syarif, A. Prugel-Bennett, and G. Wills, "Unsupervised clustering approach for network anomaly detection," in *Networked Digital Technologies. NDT 2012*, *Communications in Computer and Information Science*, R. Benlamri, Eds., vol. 293, 2012, pp. 135–145. https://doi.org/10.1007/978-3-642-30507-8_13

[22] L. A. Souto Arias, C. W. Oosterlee, and P. Cirillo, "AIDA: Analytic isolation and distance-based anomaly detection algorithm," *Pattern Recognit.*, vol. 141, p. 109607, 2023. https://doi.org/10.1016/j.patcog.2023.109607

[23] E. Panjei, L. Gruenwald, E. Leal, C. Nguyen, and S. Silvia, "A survey on outlier explanations," *VLDB J.*, vol. 31, no. 5, pp. 977–1008, 2022. https://doi.org/10.1007/s00778-021-00721-1

[24] Y. Lin and X. Li, "Back to the metrics: Exploration of distance metrics in anomaly detection," *App. Sci.*, vol. 14, p. 7016, 2024. https://doi.org/10.20944/preprints202406.0529.v2

[25] Z. Li, S. Chen, H. Dai, D. Xu, C.-K. Chu, and B. Xiao, "Abnormal traffic detection: Traffic feature extraction and DAE-GAN with efficient data augmentation," *IEEE Trans. Reliab.*, vol. 72, no. 2, pp. 498–510, 2023. https://doi.org/10.1109/TR.2022.3204349

[26] M. Said Elsayed, N.-A. Le-Khac, S. Dev, and A. D. Jurcut, "Network anomaly detection using LSTM based autoencoder," in *Proceedings of the 16th ACM Symposium on QoS and Security for Wireless and Mobile Networks*, 2020, pp. 37–45. https://doi.org/10.1145/3416013.3426457

[27] G. Bovenzi, G. Aceto, D. Ciuonzo, V. Persico, and A. Pescape, "A hierarchical hybrid intrusion detection approach in IoT Scenarios," in *GLOBECOM 2020 – 2020 IEEE Global Communications Conference*, 2020, pp. 1–7. https://doi.org/10.1109/GLOBECOM42002.2020.9348167

[28] L. I. Khalaf, B. Alhamadani, O. A. Ismael, A. A. Radhi, S. R. Ahmed, and S. Algburi, "Deep learning-based anomaly detection in network traffic for cyber threat identification," in *Proceedings of the Cognitive Models and Artificial Intelligence Conference*, 2024, pp. 303–309. https://doi.org/10.1145/3660853.3660932

[29] Eyer, "ARIMA based algorithms vs neural networks in anomaly detection," Eyer.ai, para. 1, 2024. [Online]. Available: https://eyer.ai/blog/arima-based-algorithms-vs-neural-networks-in-anomaly-detection/ [Accessed: Jan. 12, 2025].

[30] S. Xue, H. Chen, and X. Zheng, "Detection and quantification of anomalies in communication networks based on LSTM-ARIMA combined model," *Int. J. Mach. Learn. Cybern.*, vol. 13, pp. 3159–3172, 2022. https://doi.org/10.1007/s13042-022-01586-8

[31] A. Cuadra-Sanchez, J. Aracil, and J. Ramos de Santiago, "Proposal of a new information theory-based technique based on traffic anomaly detection analysis," *Int. J. Parallel Emergent Distrib. Syst.*, vol. 30, no. 6, pp. 464–477, 2015. https://doi.org/10.1080/17445760.2015.1044002

[32] C. Callegari, S. Giordano, M. Pagano, and T. Pepe, "WAVE-CUSUM: Improving CUSUM performance in network anomaly detection by means of wavelet analysis," *Comput. Secur.*, vol. 31, no. 5, pp. 727–735, 2012. https://doi.org/10.1016/j.cose.2012.05.001

[33] M. Braei and S. Wagner, "Anomaly detection in univariate time-series: A survey on the state-of-the-art," *arXiv preprint arXiv:2004.00433*, 2020. https://doi.org/10.48550/arXiv.2004.00433

[34] E. Cox, "Data analysis of Telekom dataset: Change point detection and planning," M.S. thesis, Faculty of Informatics, Eötvös Loránd University, 2017. Accessed: 2024. [Online]. Available: https://t-labs.elte.hu/wp-content/uploads/DamianoFossaThesis.pdf

[35] T. G. Smith, "pmdarima: ARIMA estimators for Python," 2017. [Online]. Available: https://alkaline-ml.com/pmdarima/ [Accessed: Nov. 10, 2024].

[36] H. Apaydin-Özkan, "Appliance-level anomaly detection by using control charts and artificial neural networks with power profiles," *Sensors*, vol. 22, no. 17, p. 6639, 2022. https://doi.org/10.3390/s22176639

[37] Z. Hasani, S. Krrabaj, and M. Krasniqi, "Proposed model for real-time anomaly detection in big IoT sensor data for smart city," *Int. J. Interact. Mob. Technol. (IJIM)*, vol. 18, no. 3, pp. 32–44, 2024. https://doi.org/10.3991/ijim.v18i03.44467

[38] H. Huang, H. Deng, J. Chen, L. Han, and W. Wang, "Automatic multi-task learning system for abnormal network traffic detection," *Int. J. Emerg. Technol. Learn. (IJET)*, vol. 13, no. 4, pp. 4–20, 2018. https://doi.org/10.3991/ijet.v13i04.8466

[39] M. M. Nadeem, Y. Raza, A. Sajid, H. Razzaq, R. Malik, and S. Vidanagamachchi, "Review analysis of web socket security: Case study," *IETI Trans. Data Anal. Forecast. (ITDAF)*, vol. 2, no. 2, pp. 56–75, 2024. https://doi.org/10.3991/itdaf.v2i2.51015

10 AUTHORS

Meriem Nabil received a B.Eng. degree in data and knowledge engineering from the School of Information Sciences in 2022. Currently pursuing doctoral studies in the field of security and artificial intelligence. Research interests encompass cybersecurity, AI, machine learning, deep learning, and IoT (E-mail: meriem.nabil@esi.ac.ma).

Meriem Hnida is a Professor at the School of Information Sciences (ESI) in Rabat, Morocco. She is a permanent member of the ITQAN research team at the LYRICA laboratory and an associate member of the "Networking, Modeling, and e-Learning (RIME)" research team at the Mohammadia School of Engineering. Her research focuses on intelligent systems in education, particularly in Intelligent

Tutoring Systems (ITS), Educational Technology, and Knowledge Engineering (E-mail: mhnida@esi.ac.ma).

Abdelhay Haqiq is a Professor at the School of Information Sciences (ESI) in Rabat, Morocco. He is a permanent member of the AL-QualSADI and ITQAN research teams at the LYRICA laboratory. His research focuses on information systems, UML profiles, multi-agent systems, and the formal specification and verification of reactive systems (E-mail: ahaqiq@esi.ac.ma).

Imane Hilal is a Professor at the School of Information Sciences (ESI) in Rabat, Morocco. She is a permanent member of the ITQAN research team at the LYRICA laboratory. Her research interests span areas such as Data Warehousing, Dependability, MDE, E-learning, Big Data, and Cloud Computing. She has been actively involved in organizing and participating in several conferences, including ISI17, MISC18, ArabWIC19, and ICSSD19 (E-mail: ihilal@esi.ac.ma).