IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

Lightweight Statistical Methods for Real-Time Anomaly Detection in Streaming Data

Sweta Dave
Lecturer, CE/IT Department
P.P. Savani University, Surat, India

Abstract: *In fields including IoT systems, cloud infrastructures, web services, industrial monitoring, and high-frequency financial markets, real-time anomaly detection is becoming more and more important. Because of their computing efficiency, interpretability, and applicability for scenarios with low latency or resources, lightweight statistical and rule-based techniques are still widely used. Moving averages, exponential smoothing, STL decomposition, seasonal hybrid ESD, and rule-based thresholding are just a few of the lightweight approaches that we carefully consider and evaluate in this paper. Their performance is evaluated using common real-world issues such as idea drift, noise, non-stationarity, multivariate dependencies, and variability in streaming environments. The study describes the benefits, drawbacks, and shortcomings of current statistical models as well as future prospects, such as hybrid statistical–machine learning frameworks meant for adaptive and scalable anomaly detection*

Keywords: Real-time anomaly detection; Statistical methods; Streaming data; IoT analytics; STL decomposition; ESD; Rule-based detection; Concept drift; Time-series analysis

I. INTRODUCTION

Real-time monitoring has become crucial due to the growing use of IoT devices, sensor networks, automated industrial systems, cloud-native apps, and high-frequency finance platforms. These systems' anomalies could be signs of malfunctions, hacking, ineffective operations, or sudden changes in behaviour. Millions of events are produced by modern apps every second, thus anomaly detection needs to be quick, understandable, and resource-efficient. Deep learning models are becoming more and more popular, however they cannot be applied in real-time situations due to their high processing costs, interpretability issues, and need for training data. However, because they offer predictable, transparent, and low latency decision-making capabilities, lightweight statistical and rule-based techniques are appropriate for mission-critical platforms, embedded systems, and edge environments.

II. BACKGROUND AND MOTIVATION

Numerous issues, including noise, missing numbers, seasonality, concept drift, and non-stationarity, arise when streaming data in practical applications. Due to latency limitations and an inability to react, traditional batch anomaly detection fails in such circumstances. Lightweight statistical and rule-based models are advised in scenarios requiring little memory utilization. Fast computation, understandable outcomes, real-time or incremental updates, and appropriateness without large labelled datasets. However, given the increasing complexity of real-world data, a careful evaluation of these methods is necessary to determine their practical benefits and drawbacks.

TABLE I: EXISTING WORK IN ANOMALY DETECTION

Method used	Purpose	Strengths	Limitations	Domains
EWMA (Exponential Weighted Moving Average)[15]	Smoothing, short-term trend detection	Very lightweight, noise-resistant	Not suitable for long-term forecasting	IoT sensors, web traffic
CUSUM / Page-Hinkley[15]	Sudden change detection	Fast, low memory, good for abrupt	Cannot detect gradual drift well	Finance, industrial monitoring

DOI: 10.48175/IJARSCT-37104

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 9, June 2026 Impact Factor: 8.2

		anomalies		
Online Z-score[8]	Point anomaly detection	Simple, very fast, incremental	Sensitive to non-stationary data	Wearables, IoT telemetry
Threshold Rules	Event triggers (above/below limit)	Easy to interpret, domain-specific	Requires manual tuning	Sensors, healthcare alerts
Statistical + Rules	Combined anomaly detection	Balances accuracy & interpretability	Still limited for complex patterns	Web traffic, finance, IoT

III. LIGHTWEIGHT STATISTICAL AND RULE-BASED METHODS

This section outlines the statistical and rule-based approaches widely used in modern monitoring systems.

3.1 Exponential Moving Average (EMA) [15] and Moving Average (MA) Moving averages smooth off abnormalities and provide a baseline for anomaly grading. EMA is preferred for real-time applications since it emphasizes current values. They are widely used in temperature sensors, financial indicators, and IT monitoring.

3.2 Holt-Winters Exponential Smoothing [4] This technique mimics level, trend, and seasonality. Because Holt-Winters smoothing can identify sporadic variations, it is useful for tracking electrical load and internet traffic.

3.3 Loess-based Seasonal-Trend Decomposition [1], or STL Using STL, a time series is divided into trend, seasonality, and residual components. It is dependable, flexible, and appropriate for data with complex seasonal trends. Remaining anomalies are found using statistical deviation.

3.4 Statistical Residual Testing [6],[3],[9] (MAD, ESD, Z-score) Residual-based methods quantify deviance and estimate typical behaviour.

IV. REAL-WORLD APPLICATIONS

Statistical and rule-based anomaly detection is used extensively across sectors:

4.1 IoT and Sensor Network Lightweight anomaly detection is crucial for battery-powered devices, factory sensors, and environmental monitoring. EMA smoothing and thresholds are commonly used due to low computational cost.

4.2 Web Traffic and Cloud Infrastructure Monitoring Web systems display daily and weekly seasonality, making STL and S-H-ESD suitable. Anomalies may include traffic spikes, latency surges, or server failures.

4.3 High-Frequency Finance Financial markets rely on EMA, Bollinger Bands, and volatility-based statistical rules for real-time fraud or crash detection.

4.4 Cybersecurity Threshold-based intrusion detection systems detect abnormal login attempts, port scans, or data exfiltration events.

V. LIMITATIONS OF STATISTICAL AND RULE-BASED METHODS

Despite their benefits, some techniques have serious drawbacks, such as strict assumptions on linearity or Gaussian noise distributions. Multivariate and nonlinear connections are not well captured. Real-world sensors are extremely noise-sensitive. gradual adaptation to notion drift. It takes a lot of manual labour to do threshold engineering. It's difficult to scale to high-dimensional data. the inability to independently learn or generalize.

VI. DISCUSSION

While statistical methods perform well in simple and stable situations, real-world systems require scalable and flexible solutions. Online machine learning, dynamic thresholding, and hybrid architectures that integrate statistical smoothing

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

with drift-aware learning models are promising methods. For example, STL residuals can be fed into an online classifier to improve detection.

VII. CONCLUSION

Because of their speed, usefulness, and comprehensibility, lightweight statistical and rule-based approaches remain essential. However, they are limited by their assumption of stationarity, vulnerability to notion drift, and incapacity to represent complex interactions. The study emphasizes the necessity of hybrid systems that combine machine learning flexibility and statistical robustness. Reducing false positives, dynamically updating baselines, and integrating multivariate context into real-time anomaly detection systems should be the main goals of future research.

REFERENCES

[1]. Cleveland, R. B., Cleveland, W. S., McRae, J. E. (1990). STL: A Seasonal-Trend Decomposition Procedure.
[2]. Chandola, V., Banerjee, A., Kumar, V. (2009). Anomaly Detection: A Survey. ACM Computing Surveys.
[3]. Ahmed, M., Mahmood, A., Hu, J. (2016). Network Anomaly Detection Techniques.
[4]. Hyndman, R.J., Athanasopoulos, G. (2018). Forecasting: Principles and Practice.
[5]. Hill, D., Minsker, B. (2010). Anomaly Detection in Streaming Sensor Data.
[6]. Laptev, N., Amizadeh, S., Flint, I. (2015). Time-Series Anomaly Detection Framework.
[7]. Hundman, K., et al. (2018). Deep Learning for Spacecraft Anomaly Detection.
[8]. Twitter Engineering Team. (2015). Seasonal Hybrid ESD Algorithm.
[9]. Xu, H., et al. (2018). Unsupervised Anomaly Detection via Forecasting.
[10]. Bontemps, L., McDermott, J., Le-Khac, N.A. (2016). Collective Anomaly Detection.
[11]. Papale, D., Black, T. (2007). Processing and Quality Control of Daytime Flux Data.
[12]. Goldstein, M., Uchida, S. (2016). A Comparative Evaluation of Unsupervised Anomaly Detection.
[13]. Gupta, M., et al. (2014). Outlier Detection for Temporal Data.
[14]. Ramaswamy, S., et al. (2000). Efficient Algorithms for Outlier Detection.
[15]. Montgomery, D. (2009). Introduction to Statistical Quality Control.