# Evaluating Temporal and Structural Anomaly Detection Paradigms for DDoS Traffic

**Yasmin Souza Lima[1], Rodrigo Moreira[1], Larissa F. Rodrigues Moreira[1], Tereza Cristina M. de B. Carvalho[3], Flávio de Oliveira Silva[2],**

[1]Institute of Exact and Technological Sciences
Federal University of Viçosa (UFV) – MG – Brazil

[2]Department of Informatics – School of Engineering
University of Minho (UMinho) – Braga – Portugal

[3]University of São Paulo (USP)
05.508-010 – São Paulo – SP – Brazil

{yasmin.lima, rodrigo, larissa.f.rodrigues}@ufv.br

terezacarvalho@usp.br, flavio@di.uminho.pt

***Abstract.*** *Unsupervised anomaly detection is widely used to detect Distributed Denial-of-Service (DDoS) attacks in cloud-native 5G networks, yet most studies assume a fixed traffic representation, either temporal or structural, without validating which feature space best matches the data. We propose a lightweight decision framework that prioritizes temporal or structural features before training, using two diagnostics: lag-1 autocorrelation of an aggregated flow signal and PCA cumulative explained variance. When the probes are inconclusive, the framework reserves a hybrid option as a future fallback rather than an empirically validated branch. Experiments on two statistically distinct datasets with Isolation Forest, One-Class SVM, and KMeans show that structural features consistently match or outperform temporal ones, with the performance gap widening as temporal dependence weakens.*

## 1. Introduction

Modern network infrastructures increasingly rely on cloud-native designs, microservice decomposition, and serverless execution models, generating large volumes of telemetry at every layer of the stack. As threat activity intensifies and Artificial Intelligence (AI)-enabled tooling accelerates adoption [Gartner, Inc. 2026], the pressing question is how to turn pervasive telemetry into a representation that makes malicious behavior separable before any detector is even selected.

The main issue is not detector choice, it is how traffic is represented before training. Most prior work commits to either temporal features or structural projections such as Principal Component Analysis (PCA) and optimizes the detector around that choice without testing paradigm fit. Recent advances span attention based sequence models for 5th Generation Mobile Network (5G) telemetry, autoencoders paired with Isolation Forest and SHAP, and clustering variants, but they still do not provide a repeatable rule to choose between temporal and structural representations in new deployments [Moreira et al. 2023].

This paper proposes a pre-training decision framework for unsupervised Distributed Denial of Service (DDoS) detection. It uses two probes, lag-1 autocorrelation of an aggregated flow signal and cumulative PCA explained variance, to prioritize temporal versus structural representations before model selection. The study evaluates only these two branches on two datasets with opposing structure, one temporally dependent and one high-dimensional 5G benchmark with weak temporal signal, using Isolation Forest (IF), One-Class Support Vector Machine (OCSVM), and KMeans. The hybrid branch is retained as a conceptual fallback for inconclusive probe combinations and is not empirically validated here. Within this scope, the results show that structural features consistently match or outperform temporal ones when autocorrelation is weak. Our contribution is therefore not a new detector, but a lightweight pre-training characterization heuristic that helps narrow the representation choice before unsupervised modelling.

## 2. Related Work

Unsupervised anomaly detection for network traffic and cloud-native systems has been addressed from multiple angles, yet no prior study explicitly compares temporal and structural feature paradigms or provides a decision rule for choosing between them. In the 5G and O-RAN domain, deep sequence models with attention have been used to detect abnormal Network Function interactions in the 3rd Generation Partnership Projec (3GPP) Service Based Architecture [Tan et al. 2025], Continuous Time Markov Chain models have targeted signaling anomalies via digital twins and safe reinforcement learning [Prince and Prabhavathi Neelakandan 2026], and LSTM autoencoders deployed in the Near-RT RIC have triggered secure slicing against malicious User Equipment [Moore et al. 2025]. For broader network traffic, IF has been extended with X-means clustering to improve detection rates [Feng et al. 2022], paired with autoencoders and SHAP-based explanations [Carrera et al. 2022], and combined with ConvLSTM for spatiotemporal modelling [Kumar et al. 2025]. Complementary efforts include calibrated one-class classification under anomaly contamination [Xu et al. 2024], dilated-convolution variational autoencoders for multi-scale patterns in mobile counters [González et al. 2024], multi-modal fusion of logs and traces in microservice architectures [Zuo et al. 2020], and a research agenda for context-aware detection in serverless environments [Nguyen et al. 2025]. Table 1 summarises the coverage of each study across ten dimensions; columns mark whether the work targets 5G/O-RAN, fuses multiple telemetry sources, employs deep learning or specific detectors (IF, OCSVM), includes explainability or mitigation, and, crucially, whether it compares temporal and structural paradigms or offers a paradigm decision rule.

**Contribution Positioning.** Unlike prior detector-specific studies, this paper addresses the paradigm choice by comparing temporal and structural representations for unsupervised DDoS detection across statistically diverse datasets. Our conceptual contribution is a lightweight decision procedure for unseen deployments using two diagnostic probes: lag-1 autocorrelation and PCA cumulative explained variance. This framework bridges detector-centric research and deployment-time decision-making, particularly in 5G environments where optimal traffic representation fluctuates.

**Table 1. Short Related Work Comparison.**

| Paper | 5G or O-RAN | Multi-source Telemetry | Deep Learning | Uses IF | Uses OCSVM | XAI | Mitigation | Temporal vs Structural | Paradigm decision rule | Open Release |
|---|---|---|---|---|---|---|---|---|---|---|
| [Zuo et al. 2020] | ○ | ● | ● | ○ | ● | ○ | ○ | ○ | ○ | ○ |
| [González et al. 2024] | ○ | ○ | ● | ○ | ○ | ○ | ○ | ○ | ○ | ● |
| [Nguyen et al. 2025] | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| [Prince and Prabhavathi Neelakandan 2026] | ● | ● | ● | ○ | ○ | ○ | ● | ○ | ○ | ○ |
| [Tan et al. 2025] | ● | ○ | ● | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| [Moore et al. 2025] | ● | ○ | ● | ○ | ○ | ○ | ● | ○ | ○ | ○ |
| [Carrera et al. 2022] | ○ | ○ | ● | ● | ○ | ● | ○ | ○ | ○ | ○ |
| [Kumar et al. 2025] | ○ | ○ | ● | ● | ○ | ● | ○ | ○ | ○ | ○ |
| [Xu et al. 2024] | ○ | ○ | ● | ○ | ○ | ○ | ○ | ○ | ○ | ● |
| [Feng et al. 2022] | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ | ○ | ○ |
| **Our Approach** | ● | ○ | ○ | ● | ● | ○ | ○ | ● | ● | ● |

## 3. Methodology

The pipeline consists of six sequential steps, illustrated in Figure 1: ❶ dataset selection and characterization, ❷ preprocessing and cleaning, ❸ feature engineering under two paradigms, ❹ unsupervised detection, ❺ evaluation against ground-truth labels, and ❻ cross-paradigm comparison.

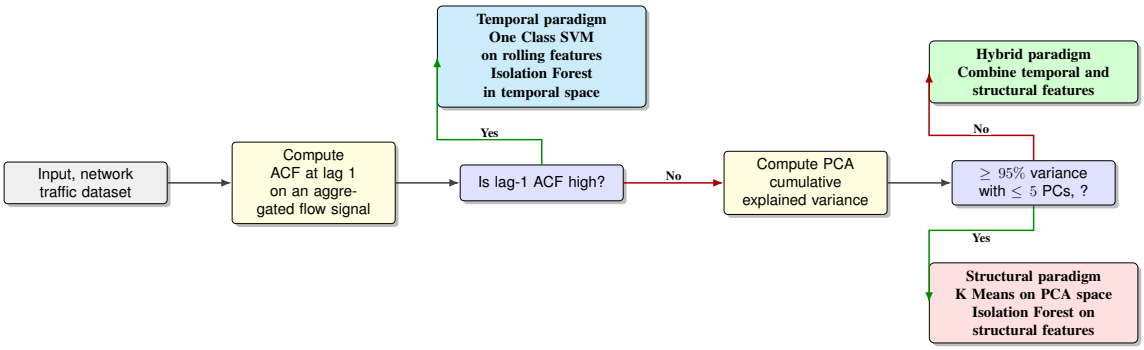


**Figure 1. Lightweight framework for prioritizing temporal or structural representations.**

### 3.1. Datasets

**Step ❶ .** Two labelled datasets with distinct traffic profiles were selected to evaluate whether detection paradigm effectiveness depends on the statistical nature of the traffic.

**CICDDoS2019.** The CICDDoS2019 benchmark [Sharafaldin et al. 2019] contains bidirectional flow records captured during simulated DDoS campaigns. Each record includes 80 features extracted by CICFlowMeter, covering packet counts, byte volumes, inter-arrival times, and flag distributions. The dataset includes eight reflection-based attack types, each stored as a separate CSV file. A stratified sample of 30 000 flows per file was drawn, yielding approximately 240 000 records with a heavily imbalanced class distribution (benign flows constitute less than 0.4%).

**5GAD.** The 5G Attack Detection dataset (5GAD) [Coldwell et al. 2022] provides simulated traffic from a 5G network slicing testbed. Records are fixed-length feature vectors of 1 024 dimensions stored as NumPy arrays. Labels are binary (`normal`/`attack`), evenly distributed across 48 320 samples. The original tensor shape was flattened to a standard two-dimensional matrix prior to analysis.

### 3.2. Preprocessing

**Step ❷ .** All experiments ran on an Intel Core i7 workstation with 32 GB RAM using Python 3.12 and scikit-learn 1.4. Numeric precision was reduced to `float32` and the OCSVM prediction phase was batched in chunks of 20 000 samples to fit memory.

Both datasets underwent a uniform cleaning procedure: non-numeric columns were removed from CICDDoS2019, infinite values and `NaN` entries were replaced with zero, and all features were standardised to zero mean and unit variance via z-score normalisation. This ensures that distance-based algorithms are not biased by scale differences across features.

### 3.3. Feature Engineering

**Step ❸ .** Two independent feature spaces were constructed from the same standardised input.

**Temporal feature space.** Rolling statistics were computed over windows of size $w \in \{10, 30, 100\}$ along the sample-index axis, preserving the sequential ordering of the original capture. Six statistics were extracted per window per base signal (rolling mean, standard deviation, maximum, minimum, first-order difference, and coefficient of variation), applied to the $\ell_2$ norm and the first five principal components, yielding $3 \times 6 \times 6 = 108$ temporal features. The coefficient of variation is defined as Eq. 1:

$$\mathrm{CV}_{w,i} = \begin{cases} \dfrac{\hat{\sigma}_{w,i}}{|\hat{\mu}_{w,i}|} & \text{if } |\hat{\mu}_{w,i}| > \epsilon, \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

where $\hat{\mu}_{w,i}$ and $\hat{\sigma}_{w,i}$ are the rolling mean and standard deviation within the window centred at sample $i$, and $\epsilon = 10^{-8}$.

**Structural feature space.** PCA was applied to the standardised data, retaining the first $d$=10 components. Each sample is projected independently onto the leading eigenvectors of the covariance matrix (Eq. 2):

$$\mathbf{z}_i = \mathbf{W}^\top \tilde{\mathbf{x}}_i, \quad \mathbf{W} \in \mathbb{R}^{p \times d}, \tag{2}$$

embedding no temporal context.

### 3.4. Unsupervised Detection Methods

**Step ❹ .** We evaluated eight experimental configurations across two datasets. Isolation Forest was applied in both temporal and structural spaces, One Class SVM was evaluated in the temporal space, and KMeans was evaluated in the structural space.

**Isolation Forest (IF).** IF isolates anomalies via recursive random partitioning; samples requiring fewer splits receive higher anomaly scores. For numerical stability under extreme class imbalance, the contamination parameter was set as $c = \min(\max(\hat{r},\, 0.01),\, 0.35)$, where $\hat{r}$ is the dataset-level attack ratio. This uses label information available only in the offline benchmark and should therefore be interpreted as evaluation-time calibration rather than a deployment-ready setting.

**One-Class SVM (OCSVM).** OCSVM learns a boundary around normal traffic in kernel space. An RBF kernel with $\gamma$=scale was used, trained on the first 5 000 samples of

each dataset under a cold-start assumption, with $\nu = \min(c,\, 0.3)$. This protocol assumes that the initial traffic segment is a reasonable proxy for nominal behaviour, which is a pragmatic benchmark choice rather than a universally valid operational assumption.

**KMeans.** KMeans partitions the feature space into $K{=}2$ clusters, consistent with the dominant binary regime and the peak Silhouette behaviour observed in Figure 5. After convergence, each cluster is associated post hoc with the majority ground-truth label among its members for external evaluation only; the clustering step itself remains fully unsupervised.

### 3.5. Evaluation and Cross-Paradigm Comparison

**Steps ❺ –❻ .** Detection performance was assessed using Precision ($P$), Recall ($R$), and F1-Score against ground-truth labels, with the attack class as the positive class. For KMeans, clustering quality was additionally measured by the Silhouette Score.

The paradigm gap for each metric $m \in \{P, R, F1\}$ quantifies the difference between the best temporal and the best structural method (Eq. 3).

$$\Delta_m = \max_{t \in \mathcal{T}} m_t - \max_{s \in \mathcal{S}} m_s, \tag{3}$$

where $\mathcal{T}$ and $\mathcal{S}$ are the temporal and structural method sets. Negative values indicate structural superiority. For unseen deployments, we interpret two diagnostic probes, namely lag-1 autocorrelation of the aggregated flow signal and cumulative variance explained by the first five principal components, as a lightweight characterization heuristic rather than a fully validated decision rule. In this paper, these probes are used primarily as a lightweight characterization heuristic to contrast strong versus weak temporal dependence and compact versus diffuse structural geometry.

## 4. Results and Discussion

We analyse how temporal dependence and feature space geometry shape the effectiveness of unsupervised DDoS detection. We compare temporal and structural pipelines on CICDDoS2019 and 5GAD, using the figures to extract four empirical insights.

**Temporal dependence separates when temporal features can help.** Figure 2 shows that CICDDoS2019 preserves sequential structure across lags, while 5GAD behaves as near independent samples beyond lag zero. Figure 3 reinforces this contrast, CICDDoS2019 is more compressible under PCA, while 5GAD spreads variance across many directions. Together, these probes indicate that temporal rollups can encode meaningful context in CICDDoS2019, whereas 5GAD offers little sequential signal to exploit.

**Both datasets contain spatial structure, but separability differs.** In Figure 4, CICDDoS2019 shows stronger overlap between classes in the leading projection, suggesting that proximity alone is not a reliable discriminator, while 5GAD exhibits clearer separation patterns that clustering can exploit. Figure 5 is consistent with this picture, low $K$ captures the dominant regimes, and 5GAD retains moderate structure for additional partitions, indicating sub profiles in attack traffic.

**Structural KMeans is the strongest overall, but temporal methods can be competitive when autocorrelation exists.** Figure 6 shows that KMeans in the structural

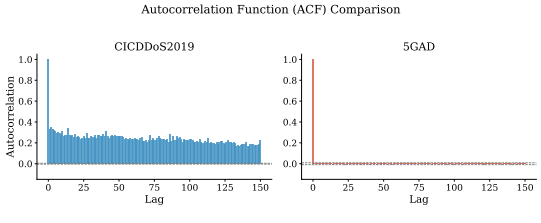


**Figure 2. ACF of the aggregated flow signal for CICDDoS2019 (left) and 5GAD (right).**

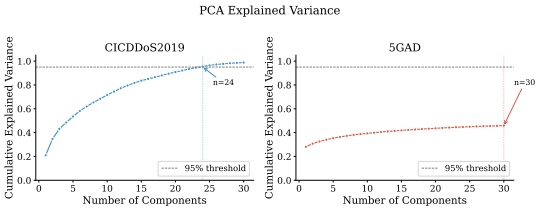


**Figure 3. Cumulative PCA explained variance for both datasets.**

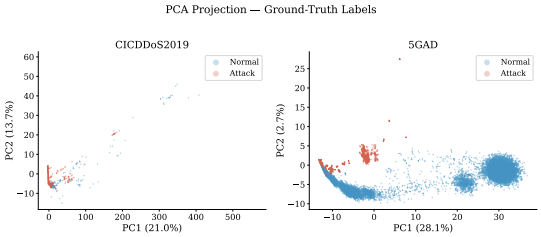


**Figure 4. PCA projection coloured by ground truth labels.**

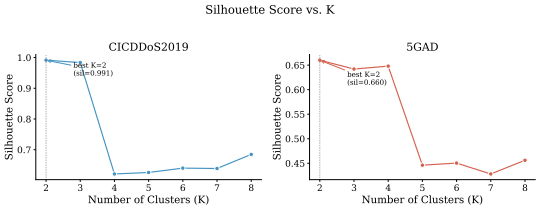


**Figure 5. Silhouette Score as a function of $K$ for both datasets.**

space achieves the best balance across metrics on both datasets. On CICDDoS2019, the best temporal method remains close, which aligns with the strong sequential dependence in Figure 2. On 5GAD, temporal methods lose effectiveness, matching the lack of exploitable temporal structure. Figure 7 summarises this shift, temporal polygons shrink on 5GAD, while the structural method retains coverage.

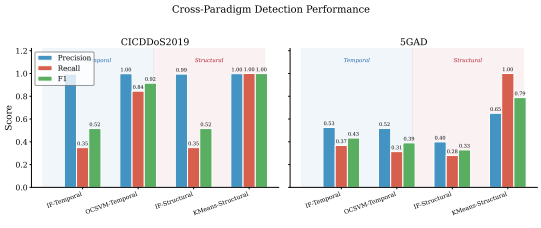


**Figure 6. Detection performance for temporal and structural methods.**

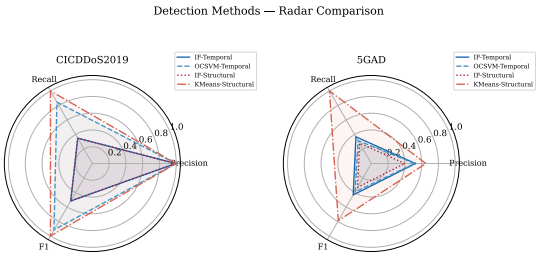


**Figure 7. Radar comparison of detection methods on both datasets.**

**Structural methods never underperform the best temporal alternative, and their advantage grows when temporal structure vanishes.** Figure 8 aggregates the paradigm gap and shows that the best structural configuration dominates across metrics on both datasets. The margin is smallest when autocorrelation is strong and temporal rollups are informative, and it increases when samples are effectively independent. This supports the characterization logic embedded in the framework. An Autocorrelation Function (ACF) probe helps indicate whether temporal context is likely to be informative, while PCA compressibility helps assess whether a compact structural projection may suffice.

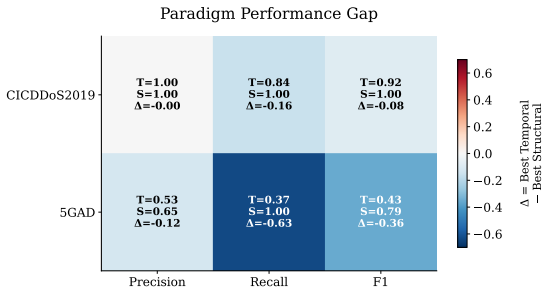


**Figure 8. Paradigm performance gap, best temporal minus best structural.**

| Method | Prec. | Rec. | F1 | Time |
|---|---|---|---|---|
| **Dataset: CICDDoS2019** | | | | |
| KMeans-Str. | 0.998 | 1.000 | 0.999 | 1.24 |
| OCSVM-Temp. | 0.998 | 0.845 | 0.915 | 26.30 |
| IF-Temp. | 0.995 | 0.349 | 0.517 | 2.34 |
| IF-Str. | 0.995 | 0.349 | 0.517 | 1.65 |
| **Dataset: 5GAD** | | | | |
| KMeans-Str. | 0.651 | 1.000 | 0.788 | 0.96 |
| IF-Temp. | 0.526 | 0.368 | 0.433 | 0.77 |
| OCSVM-Temp. | 0.518 | 0.314 | 0.391 | 5.95 |
| IF-Str. | 0.399 | 0.279 | 0.329 | 0.72 |

**Table 2. Cross-paradigm detection results.**

**Operationally, structural pipelines provide the best cost-benefit tradeoff.** Table 2 shows that structural KMeans is consistently efficient and highly effective. Temporal OCSVM can also perform well when traffic has strong temporal redundancy, but its much higher runtime limits real-time use. Isolation Forest is lightweight in both spaces, yet it is less reliable in the most imbalanced setting, so it is better used as a fast baseline than as the main option when structural clustering is available. Two cautions apply. Near-perfect results on CICDDoS2019 likely reflect strong benchmark separability and should not be directly generalized to operational traffic. Also, the IF contamination setting, the cold-start training slice for OCSVM, and the post hoc majority-label mapping for KMeans were used to stabilize offline benchmarking, not to define a fully label-free deployment recipe.

## 5. Concluding Remarks

We treat unsupervised DDoS detection as a paradigm selection problem. Structural features, KMeans on PCA, match or beat temporal alternatives, especially when autocorrelation is weak. Temporal OCSVM helps only when autocorrelation is strong. We therefore suggest a two-probe characterization heuristic based on lag-1 autocorrelation and PCA explained variance to guide representation choice before training. The evaluation relied on labelled benchmarks with known class structure, and the hybrid branch of the decision framework was not empirically validated. Extending the analysis to unlabelled production traces, calibrating the autocorrelation threshold across broader network environments, and integrating the procedure into an online pipeline that re-evaluates paradigm fit under distribution shift are natural next steps.

## Acknowledgment

The authors thank FAPEMIG (Grant #APQ00923-24), FAPESP MCTIC/CGI Research project 2018/23097-3 - SFI2 - Slicing Future Internet Infrastructures, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 for supporting this work.

## References

Carrera, F., Dentamaro, V., Galantucci, S., Iannacone, A., Impedovo, D., and Pirlo, G. (2022). Combining unsupervised approaches for near real-time network traffic anomaly detection. *Applied Sciences*, 12(3).

Coldwell, C., Conger, D., Goodell, E., Jacobson, B., Petersen, B., Spencer, D., Anderson, M., and Sgambati, M. (2022). Machine learning 5g attack detection in programmable logic. In *2022 IEEE Globecom Workshops (GC Wkshps)*, pages 1365–1370.

Feng, Y., Cai, W., Yue, H., Xu, J., Lin, Y., Chen, J., and Hu, Z. (2022). An improved X-means and isolation forest based methodology for network traffic anomaly detection. *PLOS ONE*, 17(1):1–18.

Gartner, Inc. (2026). Gartner forecasts worldwide end-user spending on information security to total $240 billion in 2026.

González, G. G., Tagliafico, S. M., Fernández, A., Sena, G. G., Acuña, J., and Casas, P. (2024). 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*, 21(2):1601–1616.

Kumar, A., Kumar, A., Raja, R., Dewangan, A. K., Kumar, M., Soni, A., Agarwal, D., and Saudagar, A. K. J. (2025). Revolutionising anomaly detection: a hybrid framework for anomaly detection integrating isolation forest, autoencoder, and conv. lstm. *Knowledge and Information Systems*, 67(12):11903–11953.

Moore, J., Abdalla, A. S., Reshi, Z., and Marojevic, V. (2025). Anomaly detection and mitigation in o-ran networks using an lstm-rnn autoencoder and secure slicing. In *MILCOM 2025 - 2025 IEEE Military Communications Conference (MILCOM)*, pages 1–6.

Moreira, R., Rodrigues Moreira, L. F., and de Oliveira Silva, F. (2023). An intelligent network monitoring approach for online classification of Darknet traffic. *Computers and Electrical Engineering*, 110:108852.

Nguyen, C., Elmroth, E., and Bhuyan, M. (2025). Silent failures in stateless systems: Rethinking anomaly detection for serverless computing. In *2025 IEEE International Conference on Service-Oriented System Engineering (SOSE)*, pages 8–19.

Prince, G. and Prabhavathi Neelakandan, R. (2026). Ai-driven analysis and mitigation of control-plane signaling anomalies in next-generation mobile networks. *IEEE Access*, 14:11129–11148.

Sharafaldin, I., Lashkari, A. H., Hakak, S., and Ghorbani, A. A. (2019). Developing realistic distributed denial of service (ddos) attack dataset and taxonomy. In *2019 International Carnahan Conference on Security Technology (ICCST)*, pages 1–8.

Tan, Y., Liu, J., Li, Y., and Wang, J. (2025). Deep learning-based proactive anomaly detection for 5g core control plane network function interactions. *IEEE Transactions on Cognitive Communications and Networking*, 11(6):4210–4222.

Xu, H., Wang, Y., Jian, S., Liao, Q., Wang, Y., and Pang, G. (2024). Calibrated one-class classification for unsupervised time series anomaly detection. *IEEE Transactions on Knowledge and Data Engineering*, 36(11):5723–5736.

Zuo, Y., Wu, Y., Min, G., Huang, C., and Pei, K. (2020). An intelligent anomaly detection scheme for micro-services architectures with temporal and spatial data analysis. *IEEE Transactions on Cognitive Communications and Networking*, 6(2):548–561.