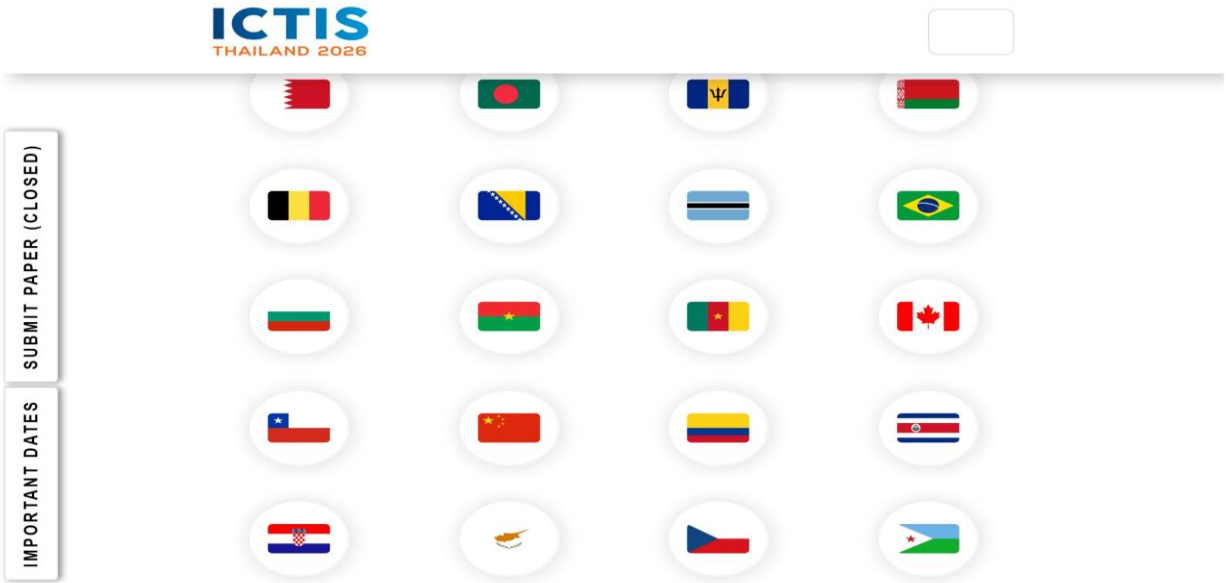

SUBMIT PAPER (CLOSED)
IMPORTANT DATES
ICTIS
THAILAND 2026
11th INTERNATIONAL CONFERENCE ON
INFORMATION AND COMMUNICATION TECHNOLOGY
FOR INTELLIGENT SYSTEMS
9 - 11 April, 2026
Bangkok, Thailand
The 11th Edition of the conference will be in Hybrid Mode [10 - April 2026 - Physical Mode || 9 - 10 - 11 April 2026 - Digital Mode]


SUBMIT PAPER (CLOSED)
IMPORTANT DATES
ICTIS
THAILAND 2026
11th INTERNATIONAL CONFERENCE ON
INFORMATION AND
COMMUNICATION TECHNOLOGY
FOR INTELLIGENT SYSTEMS
9 - 11 April, 2026 | Bangkok, Thailand
The 11th Edition of the conference will be in Hybrid Mode [10 - April 2026 - Physical Mode || 9 - 10 - 11 April 2026 - Digital Mode]


ICTIS
THAILAND 2026
SUBMIT PAPER (CLOSED)
IMPORTANT DATES

WELCOME TO ICTIS 2026

ICTIS Conference is now in the 12th year of its commencement and as a technical celebration to attract researchers from across the world, this year ICTIS 2026 conference will be held in 2 phases. The 1st phase of ICTIS 2026 will be held in Bangkok, Thailand during April 9-11, 2026 and the 2nd phase will be held in New York, USA during 22-23 May 2026.

On behalf of the Organizing Committee of **ICTIS 2026**, we take the pleasure to invite you to The **11th International Conference on Information and Communication Technology for Intelligent Systems** that will be held on **APRIL 9 - 11, 2026 in Bangkok, Thailand.**

This conference will provide the participants with opportunities to discuss and explore areas related to the Theory, Development, Applications, Experiences and Evaluation of Interaction Sciences with fellow students, researchers and practitioners. Conference may concern any topic within the conference scope. Workshops may be related to any topics within the conference scope. The conference is devoted to increase the understanding role of Information and Communication Technology, how Intelligent Systems has day by day evolved. The conference will provide a platform for bringing forth significant research and literature across the field of Information and Communication Technology for Intelligent Systems and provide an overview of the technologies waiting unveiling. This interaction will be the focal point for leading experts to share their insights, provide guidance and address participant's questions and concerns.

The advent of the World Wide Web has sparked renewed interest in the area of intelligent information technologies. There is a growing interest in developing intelligent technologies that enable users to accomplish complex tasks in web-centric environments with relative ease, utilizing such technologies as intelligent agents, distributed computing in heterogeneous environments, and computer supported collaborative work. The mission of the International Conference on ICT for Intelligent Systems is to bring together researchers in related fields such as information systems, distributed AI, intelligent agents, and collaborative work, to explore and discuss various aspects of design and development of intelligent technologies. This journal provides a forum for academics and practitioners to explore research issues related to not only the design, implementation and deployment of intelligent systems and technologies, but also economic issues and organizational impact.

ICTIS Thailand 2026 : Digital Session Reminder || Paper ID: 169

External

Inbox

ICTIS 2026 Thailand 8 Apr 2026, 06:25

Dear Michael David,

Paper ID: 169

Title: Conf-Gate XGBoost-RF Hybrid Model for Multi-Class Anomaly Classification in 5G-Enabled eMTC IoT Networks

Greetings to you from Team ICTIS 2026!

On 09th to 11th April the ICTIS 2026 Digitally session begins at 09:30 AM through the Zoom. All Mention Times are In Bangkok Time.

The complete agenda of the conference including all technical sessions is available at: https://ictisthailand.com/agenda.html

If you are not able to get the ZOOM Links through the agenda you can check this page for convenience: https://ictisthailand.com/zoom.html

Virtual Room A - https://gr-foundation.zoom.us/j/84907847047
Meeting ID : 849 0784 7047
Passcode : 123456

Virtual Room B
- https://us06web.zoom.us/j/87813660101?pwd=bjFvRKhXUjHD8ax1t7aae1m8SkmmJ3.1

Meeting ID : 878 1366 0101
Passcode : 123456

Virtual Room C
- https://us06web.zoom.us/j/85850005625?pwd=yK2XKB97JQuqUVtWIcdLLCnZggLqlN.1
Meeting ID : 858 5000 5625
Passcode : 123456

Virtual Room D
- https://us06web.zoom.us/j/82998915684?pwd=sSVCUZdtJaXjB4UnRP7Hba4T5uCFUx.1
Meeting ID : 829 9891 5684
Passcode : 123456
Virtual Room E
- https://us06web.zoom.us/j/82800564767?pwd=a1mRdUK7477ebfK2E0vlnRy68acTW3.1
Meeting ID : 828 0056 4767
Passcode : 123456

Virtual Room F
- https://us06web.zoom.us/j/86156538061?pwd=sHWNVoyDZkX3JDMt7Yy3TCMNzt29m9.1
Meeting ID : 861 5653 8061
Passcode : 123456

Virtual Room G
- https://us06web.zoom.us/j/81361498003?pwd=Xaiame6k6lstUwiVwtyqhWVtxjmyFz.1
Meeting ID : 813 6149 8003
Passcode : 123456

We look forward to the presence of all authors and co-authors for the sessions.

If you have any queries during your Digital Technical Session kindly write it down to the Host in the Zoom chat box.

Best regards,
Team ICTIS 2026
Bangkok Thailand

REVIEWERS COMMENTS

Review 1

Overall evaluation

2: accept

- This manuscript presents a two-stage hybrid anomaly detection model, termed Conf-Gate XGBoost-RF, to address class imbalance in multi-class classification tasks in 5G-enabled eMTC IoT networks.

-The model uses a confidence-gating mechanism to combine the high efficiency of XGBoost with the robustness of Random Forests. The evaluation on the CICIoT2023 dataset demonstrates significant improvement in rare-class detection while maintaining high efficiency. The paper is well-written, addresses a critical issue in IoT security, and contributes a methodologically sound, operationally viable solution.

-The confidence-based routing to a specialist model is well-conceived and justified. Further elaboration on how the threshold value θ is chosen and how sensitive the performance is to its variation would improve reproducibility.

-The study effectively reports metrics like Macro F1 and per-class recall, which is commendable. Including confusion matrices for a few rare classes (e.g., CommandInjection, SqlInjection) would further highlight the effectiveness of the proposed method.

-The inclusion of multiple baselines such as XGB-CTGAN, XGB-GAN, and XGB-RF-SMOTE demonstrates rigor. However, a concise table summarizing the configuration and main takeaway from each baseline (perhaps in the appendix) would improve clarity.

-It would be beneficial to provide a visual summary (e.g., flowchart or block diagram) of the proposed model on an early page to guide readers through the two-stage architecture.

Review 2

Overall evaluation

2: accept

- The authors propose a hybrid machine learning architecture combining XGBoost and Random Forests, empowered by a confidence gating mechanism, to enhance anomaly detection in imbalanced 5G eMTC IoT datasets.

-The model is rigorously tested and performs well, particularly in rare-class recall, which is often neglected. The paper provides methodological depth, practical deployment insight, and a strong experimental foundation. It is a meaningful contribution to the field of IoT security and anomaly detection.

-The reported inference time (~0.189 µs/sample) is impressive. Including a deployment use case or a simulation of real-time traffic would strongly validate the practical viability of the method.

-The selection of Random Forest for the specialist model is reasonable. However, a brief justification or empirical comparison (even if informal) with a neural-based alternative like a lightweight MLP would enrich the study.

-The class taxonomy (Figure 2, Page 8) is well-structured. A short explanation on how this grouping supports model interpretability or facilitates generalization would enhance its utility.

-Some mathematical expressions (e.g., Equations 11, 20, 24) are clear but could benefit from a one-line intuitive explanation alongside.

It would be helpful to highlight which attack classes belong to the low-volume set directly within Table 2 rather than referring to text later.

ICTIS
THAILAND 2026

11th Edition ICTIS 2026

CERTIFICATE

THIS IS TO CERTIFY THAT

MICHAEL DAVID

HAS DIGITALLY PRESENTED A PAPER TITLED

CONF-GATE XGBOOST-RF HYBRID MODEL FOR MULTI-CLASS ANOMALY CLASSIFICATION IN 5G-ENABLED EMTC IOT NET-WORKS

AT THE

11th INTERNATIONAL CONFERENCE ON INFORMATION AND COMMUNICATION TECHNOLOGY FOR INTELLIGENT SYSTEMS

NILANJAN DEY, Ph.D
TPC CHAIR

TACHANUN KANGWANTRAKOOL, Ph.D
LOCAL CONFERENCE CHAIR

AMIT JOSHI, Ph.D
INTERNATIONAL CONFERENCE CHAIR

9 - 11 APRIL 2026 | BANGKOK, THAILAND

DIGITAL PLATFORM : ZOOM

Local Organizing Partner: IADDA International Auditors for Digital and Data Management Association

International Organizing Partner: GR FOUNDATION

International Managing Partner: GR SCHOLASTIC

Publication Partner: SPRINGER NATURE

Global Knowledge Research Foundation

Dear ***Michael David,***

Greetings to You…!

On behalf of the 11th International Conference on Information and Communication Technology for Intelligent Systems (ICTIS 2026), Global Knowledge Research Foundation and G R Scholastic, we thank you for being a part of the conference held in hybrid mode. We also thank you for your graceful presence during the digital conference on 9th - 11th April, 2026 held through ZOOM. It was sheer joy to have you at the International Conference and further listen to your knowledgeful address.

We hope to see you again for the next International Conference on Information and Communication Technology for Intelligent Systems (ICTIS 2027) at Bangkok, Thailand.

We treasure your contributions and time with us.

On behalf of Team ICTIS 2026,

With Best Regards,

***Amit Joshi**, PhD*
International Conference Chair - ICTIS 2026
Director - Global Knowledge Research Foundation,
Email - amit@gr.foundation

G R SCHOLASTIC LLP

ICTIS THAILAND 2026

Receipt No :	3804

RECEIPT

Date :	30-03-2026

Received with thanks the sum of USD: Five Hundred And Fourty Point Eight

Mode of Payment : Online payment

ICTIS Thailand Ref. No : 202526169

M/S : MICHAEL DAVID

USD	$ 540.8/-

11th International Conference on Information and Communication Technology for Intelligent Systems
Date : April 09 - 11, 2026
Conference Venue : Bangkok Marriott Hotel Sukhumvit, Sukhumvit Thonglor, 2, Sukhumvit Road Soi 57, Klongtan-Nua, Wattana, Bangkok 10110, THAILAND.

For, G R SCHOLASTIC LLP

G R SCHOLASTIC LLP – G/R Scholastic

Authorized Signature

ICTIS THAILAND 2026

11th Edition ICTIS 2026

CERTIFICATE

This is to certify that

Michael David, Chekwas Ifeanyi Chikezie, Abraham Usman Usman, Sulieman Zubair, Henry Ohiani Ohize, Joseph Ojeniyi

has contributed a paper titled

Conf-Gate XGBoost-RF Hybrid Model for Multi-Class Anomaly Classification in 5G-Enabled eMTC IoT Networks

in the

11th International Conference on
Information and Communication Technology for Intelligent Systems

The paper has also been selected for publication in the (ICTIS) Conference as per fulfilment of guidelines issued by Springer.

NILANJAN DEY, Ph.D	TACHANUN KANGWANTRAKOOL, Ph.D	AMIT JOSHI, Ph.D
TPC CHAIR	LOCAL CONFERENCE CHAIR	INTERNATIONAL CONFERENCE CHAIR

9 - 11 APRIL 2026 | BANGKOK, THAILAND — DIGITAL PLATFORM : ZOOM

Local Organizing Partner	International Organizing Partner	International Managing Partner	Publication Partner
IADDA	GR FOUNDATION	GR SCHOLASTIC	SPRINGER NATURE

Conf-Gate XGBoost-RF Hybrid Model for Multi-Class Anomaly Classification in 5G-Enabled eMTC IoT Networks

Michael David [1][0000-0003-2879-225X], Chekwas Ifeanyi Chikezie [2][0000-0001-9870-3318], Abraham Usman Usman [3][0000-0002-3154-5156], Sulieman Zubair [4][0000-0001-5242-3820], Henry Ohiani Ohize [5][0000-0001-5698-1889], and Joseph Ojeniyi [6][0000-0003-0707-4152]

[1,2,3,4] Department of Telecommunications Engineering,
[5] Department of Electrical Engineering,
[6] Department of Cyber Security Science,
Federal University of Technology,
Minna, Niger State, Nigeria
chekwas.pg2010285@st.futminna.edu.ng

Abstract. The rapid growth of Internet of Things (IoT) deployments in 5G Enhanced Machine-Type Communication (eMTC) networks has significantly increased the network attack surface. A major challenge for Network Anomaly Detection Systems (NADS) in this environment is severe class imbalance, where dominant benign traffic obscures rare but high-impact attacks, leading to poor minority-class detection. This paper presents Conf-Gate XGBoost-RF, a two-stage hybrid anomaly detection architecture designed to address this limitation without compromising real-time performance. The framework employs a high-speed XGBoost classifier for initial screening and a confidence-gated mechanism that selectively routes low-confidence predictions to a specialist Random Forest trained on synthetically balanced data. Evaluation on the large-scale CICIoT2023 dataset shows that the proposed model achieves 99.32% accuracy and a Macro F1-score of 0.80, substantially outperforming single-stage baselines. Notably, recall for critical low-volume attacks, such as Command Injection, improves by over 34%. With an average inference latency of 0.87 ms, the proposed approach remains compatible with the stringent low-latency requirements of 5G eMTC control signaling, demonstrating a practical balance between computational efficiency and rare-attack sensitivity.

Keywords: 5G eMTC, Network Anomaly Detection, Class Imbalance, Hybrid Learning, IoT Security.

1 Introduction

The rapid expansion of the Internet of Things (IoT) has reshaped modern infrastructure by interconnecting billions of devices across domains such as smart cities and industrial automation [1, 2]. This connectivity is strengthened by 5G networks, which provide high data rates, ultra-low latency, and massive device support [3]. Within this ecosystem, Enhanced Machine-Type Communication (eMTC, LTE Cat-M1) serves as a key Low-Power Wide-Area technology introduced in 3GPP Release 13 to support low-complexity IoT devices over existing LTE infrastructure [4]. Compared with NB-IoT,

eMTC offers higher data rates of up to 1 Mbps, lower latency in the range of 10 to 15 ms, mobility support, and extended coverage, making it suitable for large-scale IoT deployments.

Despite these advantages, the convergence of 5G and eMTC introduces significant security challenges. Device heterogeneity and the resource-constrained nature of eMTC nodes limit the adoption of strong security mechanisms, leaving many devices vulnerable to compromise [5]. Empirical studies indicate that exposed IoT devices can be discovered and exploited within minutes, enabling attacks such as distributed denial-of-service amplification, data exfiltration, and unauthorized access [5]. At 5G scale, the massive number of connected devices amplifies the impact of such compromises.

Network Anomaly Detection Systems (NADS) are therefore essential for securing 5G-enabled eMTC environments. However, their deployment is complicated by severe class imbalance in real-world traffic, where benign behavior and common attacks dominate, while sophisticated and high-impact intrusions occur infrequently. Conventional machine learning models, including ensemble approaches that aggregate predictions from all classifiers simultaneously, tend to favor majority classes. As a result, they achieve high overall accuracy but exhibit poor sensitivity to rare yet critical threats.

To address this limitation, this work proposes Conf-Gate, a confidence-gated hybrid anomaly detection architecture that combines the computational efficiency of XGBoost with the robustness of Random Forests. Unlike traditional ensemble methods, Conf-Gate adopts a sequential gating mechanism. Incoming traffic is first processed by a high-speed XGBoost classifier, and predictions with high confidence are accepted directly. Samples associated with low confidence are selectively routed to a specialist Random Forest trained on balanced data, enabling more discriminative analysis of rare and ambiguous attack patterns.

To clarify the operational workflow of the proposed approach, Figure 1 illustrates the Confidence-Gated (Conf-Gate) hybrid architecture, highlighting the routing of samples between the global and specialist classifiers based on prediction confidence.

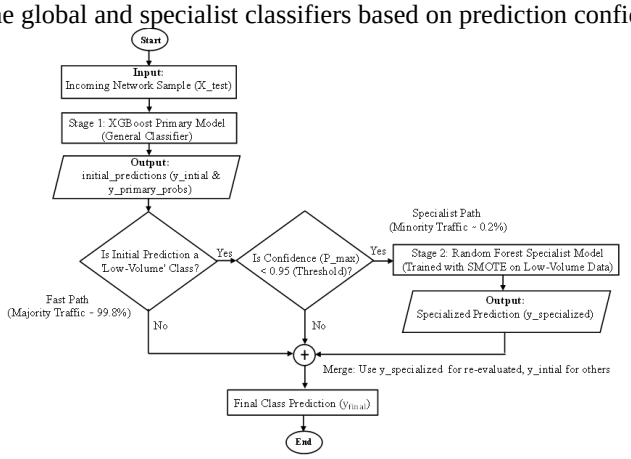


Figure 1: Confidence-Gated (Conf-Gate) Hybrid Model Architecture Flowchart

As shown in Figure 1, the two-stage design balances detection sensitivity and computational efficiency by reserving intensive analysis for only uncertain samples. The

main contributions of this work are threefold: (i) the definition of a hierarchical taxonomy for 5G IoT attacks to enhance interpretability; (ii) the introduction of a confidence-gated hybrid detection framework that jointly optimizes inference speed and rare-class sensitivity; and (iii) comprehensive validation on the CICIoT2023 dataset, demonstrating improved minority attack detection over standalone classifiers and oversampling-based methods while maintaining sub-millisecond latency compatible with eMTC requirements.

2 Related Work

The Network Anomaly Detection Systems (NADS) for IoT have progressively shifted from signature-based techniques to Machine Learning (ML) and Deep Learning (DL) approaches capable of detecting zero-day attacks. Recent studies have examined these methods within 5G environments, where scalability and latency constraints are more pronounced. Chikezie et al. [6] highlighted the scalability challenges specific to 5G eMTC networks, showing that conventional security appliances struggle to cope with the massive connection density of modern IoT deployments. Similarly, Reis [7] proposed an AI-driven anomaly detection framework for 5G-enabled smart cities and demonstrated that ML models can accurately classify common attack vectors. , Fares et al. [8] introduced a hybrid Swin Transformer and LSTM architecture to capture temporal traffic dependencies. Although effective, such deep architectures incur high computational cost and latency, limiting their suitability for resource-constrained eMTC scenarios.

A fundamental challenge in IoT anomaly detection is severe class imbalance, where malicious traffic represents only a small fraction of observed data. The CICIoT2023 dataset released by Neto et al. [9] exposes this imbalance, as real-world IoT traffic is overwhelmingly benign. Chikezie et al. [3] further showed that in high-volume 5G environments, standard classifiers often report high accuracy while failing to detect rare but high-impact attacks. To address this, data-level augmentation techniques such as the Synthetic Minority Over-sampling Technique (SMOTE) have been widely adopted. However, oversampling can introduce noise and overfitting, and repeated resampling increases computational overhead. Generative Adversarial Networks (GANs) have also been explored for synthesizing attack traffic, but they require substantial training time and exhibit instability in convergence, which hinders lightweight deployment [10].

Ensemble and hybrid learning methods attempt to improve robustness by combining multiple classifiers. Conventional approaches such as voting or stacking ensembles aggregate predictions from all base learners for every input sample [11]. While this improves accuracy, it increases inference latency proportionally to ensemble size, which conflicts with the strict real-time requirements of 5G eMTC control signaling. In contrast, the approach proposed in this work introduces a confidence-gated hybrid architecture that conditionally invokes a specialist classifier only for low-confidence predictions. By decoupling computational cost from model complexity, this strategy enables effective rare-class detection while preserving the low-latency throughput required in large-scale IoT deployments.

3 Methodology

3.1 Preprocessing Pipeline

Given the severe class imbalance and large scale of IoT traffic data, a compact and computationally efficient preprocessing pipeline was designed. The CIC-IoT2023 dataset, comprising traffic from 105 IoT devices and 33 attack classes, was adopted as the benchmark. The pipeline supports binary, multi-class, and unsupervised learning paradigms.

The dataset, originally distributed across N CSV files, was consolidated into a single composite dataset as defined in Equation 1:

$$\mathcal{D}_{\text{combined}} = \bigcup_{i=1}^{N} \mathcal{D}_i \tag{1}$$

To reduce memory overhead for the ($\sim$ 12.8 GB) dataset, numerical features were downcast to the smallest lossless data types, as formalized in Equation 2:

$$\text{Minimize} \sum_{i=1}^{n} \text{Memory}\,(x_i) \;\; \text{subject to} \; x_i^{\text{original}} = x_i^{\text{downcast}} \tag{2}$$

Although the dataset contains no missing values, zero-imputation was retained for generalizability. Non-informative fields (timestamps, flow identifiers) were removed, and features exhibiting multicollinearity ($\rho > 0.90$) were eliminated. The full deterministic preprocessing workflow is summarized in Algorithm 1.

Algorithm 1: **Deterministic Pre-processing**

Input: Raw CSV path p**, memory cap** M**, random seed** s

Output: $\{\mathbf{X_{train}}, \mathbf{y_{train}}, \mathbf{X_{val}}, \mathbf{y_{val}}, \mathbf{X_{test}}, \mathbf{y_{test}}\}$**, scalers, encoders**

assert SchemaConsistency$(\boldsymbol{p})$

$\boldsymbol{df} \leftarrow$ pd.read_csv $(p,$ nrows $= \left[\frac{M}{2.5\ kB}\right])$

$\boldsymbol{df} \leftarrow$ DownCast $(\boldsymbol{df})$

$\boldsymbol{df} \leftarrow$ ImputeNaN $(\boldsymbol{df}, \mathbf{0})$

$\boldsymbol{df} \leftarrow$ DropDuplicates $(\boldsymbol{df})$

$\mathbf{X}, \mathbf{y} \leftarrow$ Encode $(\boldsymbol{df})$

$\mathbf{X} \leftarrow$ RemoveCorrelated $(\mathbf{X}, \mathbf{0.95})$

$(\mathbf{X_{train}}, \mathbf{y_{train}}), (\mathbf{X_{val}}, \mathbf{y_{val}}), (\mathbf{X_{test}}, \mathbf{y_{test}}) \leftarrow$ StratifiedSplit $(\mathbf{X}, \mathbf{y}, \mathbf{s})$

Persist $(\mathbf{X_{train}}, \mathbf{y_{train}}, \mathbf{X_{val}}, \mathbf{y_{val}}, \mathbf{X_{test}}, \mathbf{y_{test}}$**, scalers)**

After preprocessing, dimensionality reduction was applied to avoid unnecessary computation in resource-constrained 5G-eMTC environments. The final 35 retained features used for training are listed in Table 1.

3.2 Feature Importance and Class Taxonomy

To justify feature selection, a Random-Forest-based importance analysis was conducted using Gini impurity reduction. As shown in Figure 2, FIN_flag_number

(0.2095) and ICMP (0.1334) emerged as the most discriminative features, consistent with their roles in DoS flooding and reconnaissance attacks. Conversely, protocol-specific indicators such as Telnet exhibited negligible importance and were excluded to reduce dimensionality.

This outcome aligns with domain knowledge: FIN flags reflect abnormal connection termination behavior, while ICMP traffic is strongly associated with scanning and probing activities.

3.3 Label Transformation and Resampling

The dataset exhibits extreme imbalance, with high-volume attacks (e.g., Mirai) dominating rare but critical threats such as SqlInjection. To mitigate this, SMOTE was applied exclusively to the training set, ensuring unbiased validation and testing.

SMOTE generates synthetic minority samples by interpolating between a minority instance x_i and its nearest neighbor x_{NN}, as defined in Equation 3:

$$x_{new} = x_i + \lambda(x_{NN} - x_i), \qquad \lambda \sim U(0,1) \tag{3}$$

This strategy enables the model to learn robust decision boundaries for rare classes without duplicating samples. The re-sampling procedure is summarized in Algorithm 2.

Algorithm 2: **Re-sampling Procedure**

Input: Original training set $\{\mathbf{X_{train}}, \mathbf{y_{train}}\}$, sampling ratio $\boldsymbol{\theta}$, neighbor count $\boldsymbol{k}$

Output: Balanced training set $\{\mathbf{X_{train^\star}}, \mathbf{y_{train^\star}}\}$

1: $(\mathbf{X_{sub}}, \mathbf{y_{sub}}) \leftarrow$ StratifiedSample $(\mathbf{X_{train}}, \mathbf{y_{train}}, \boldsymbol{\theta})$

2: eligible $\leftarrow \{ \boldsymbol{c} \mid \mathbf{Count}(\mathbf{y_{sub}}, \boldsymbol{c}) > \boldsymbol{k} \}$

3: **if** $|\mathbf{eligible}| \geq \mathbf{2}$ **then**

4: $(\mathbf{X_{train^\star}}, \mathbf{y_{train^\star}}) \leftarrow$ SMOTE $(\mathbf{X_{sub}}, \mathbf{y_{sub}}, \boldsymbol{k})$

5: **else**

6: $(\mathbf{X_{train^\star}}, \mathbf{y_{train^\star}}) \leftarrow (\mathbf{X_{sub}}, \mathbf{y_{sub}})$ **// fallback**

7: Persist $(\mathbf{X_{train^\star}}, \mathbf{y_{train^\star}}, \mathbf{X_{val}}, \mathbf{y_{val}}, \mathbf{X_{test}}, \boldsymbol{y_{test}})$

All numerical features were then standardized using Z-score normalization (Equation 4) to facilitate stable convergence:

$$x_i^{scaled} = \frac{x_i - \mu_i}{\sigma_i} \tag{4}$$

3.4 Conf-Gate Hybrid Architecture

The proposed Conf-Gate model adopts a two-stage cascade following a fast-path/slow-path design. As illustrated in Figure 1, inference is governed by the gating function in Equation 5:

$$f_{\text{Conf-Gate}}(x) = \begin{cases} f_{\text{spec}}\left(x f_{\text{prim}}(x)\right), & \text{if } \hat{y}_1 = k \in \mathcal{L} \wedge p_{\max} < \theta, \\ f_{\text{prim}}(x), & \text{otherwise,} \end{cases} \tag{5}$$

Here, $\mathcal{L}$ denotes the low-volume class set and (\theta) is the confidence threshold. This mechanism ensures that computationally expensive specialist inference is invoked only for uncertain predictions involving rare attacks.

Stage-1: Global XGBoost Learner

The primary classifier (f_{prim}) is an XGBoost model optimized for high-throughput inference. It minimizes the regularized cross-entropy loss in Equation 6:

$$\mathcal{L}_{\text{XGB}} = \text{CrossEntropyLoss}\big(\text{y}, f_{\text{XGB}}(\text{x})\big) + \Omega(\text{model}) \quad (6)$$

Hyperparameters ($n\_estimators = 300, \max\_depth = 6$) were selected via 5-fold cross-validation.

Stage-2: Specialist Random Forest

The specialist classifier $\big(f_{\text{spec}}\big)$ is a Random Forest trained solely on SMOTE-balanced low-volume classes. Random Forest was preferred over neural alternatives due to its robustness on small tabular datasets, resistance to overfitting, and predictable inference latency.

Inference Workflow

The latency-aware inference logic is summarized in Algorithm 3.

Algorithm 3: Conf-Gate Inference

Input: A data sample $\mathbf{x}$, the trained global model $\boldsymbol{f}_{\mathbf{prim}}$, the trained specialist model $\boldsymbol{f}_{\mathbf{spec}}$, the set of low-volume classes $\boldsymbol{\mathcal{L}}$, and the confidence threshold $\boldsymbol{\theta}$.

Output: A predicted class label for the sample $\mathbf{x}$.

$\hat{\mathbf{y}}_{\mathbf{1}}, \mathbf{pmax} \leftarrow \mathbf{predict}_{\mathbf{proba}}\big(f_{\mathbf{prim}}(\mathbf{x})\big)$
if $\hat{\mathbf{y}}_{\mathbf{1}} \in \boldsymbol{\mathcal{L}}$ **and** $\boldsymbol{p}_{\mathbf{max}} < \boldsymbol{\theta}$ **then**
    $\boldsymbol{z} \leftarrow \boldsymbol{f}_{\mathbf{spec}}\big(\boldsymbol{\phi}(\boldsymbol{x})\big)$
    return $\boldsymbol{z}$
else
    return $\hat{\mathbf{y}}_{\mathbf{1}}$

3.5 Evaluation Metrics

To ensure reliable assessment under imbalance, metrics beyond accuracy were employed.

Matthews Correlation Coefficient (MCC) provides a balanced measure for multi-class imbalanced problems and is defined in Equation (7):

$$\text{MCC} = \frac{\text{c}\sum_k \sum_l \sum_m \epsilon_{klm} - \sum_k \sum_l \alpha_k \beta_l}{\sqrt{(c^2 - \sum_k \alpha_k^2)\,(c^2 - \sum_l \beta_l^2)}} \quad (7)$$

Geometric Mean (G-Mean) evaluates the balance between sensitivity and specificity:

$$\text{G} - \text{Mean} = \sqrt{\text{Sensitivity} \times \text{Specificity}} = \sqrt{\text{Recall} \times \text{TNR}} \quad (8)$$

Cohen's Kappa ((\kappa)) corrects observed agreement for chance effects and is computed as:

$$\kappa = \frac{P_{\text{o}} - P_{\text{e}}}{1 - P_{\text{e}}} \quad (9)$$

4 Experiments and Results

This section evaluates the proposed Conf-Gate XGBoost-RF model against multiple baselines under severe class imbalance. The analysis proceeds from baseline performance to rare-class behavior, threshold sensitivity, ablation studies, comparison with state-of-the-art methods, and operational viability under simulated real-time traffic.

4.1 Baseline Model Performance

To establish a competitive reference, six baseline models were evaluated, including single classifiers (XGBoost and Random Forest) and cascade variants incorporating oversampling and generative augmentation. Table 2 summarizes their overall accuracy, weighted F1-score, computational cost, and model size, providing insight into both detection performance and operational overhead.

Table 2: Comparative Performance of Baseline Models on the Test Set.

Model	Accuracy	F1-Score (Weighted)	MCC	Training Time (s)	Inference Time (ms/sample)	Model Size (MB)
B1: XGB-GAN	0.9933	0.9931	0.9926	1887.66	0.0003	27.04
B2: XGB-CTGAN	0.9933	0.9931	0.9926	3197.94	0.0004	27.04
B3: XGB-RF-SMOTE	0.9933	0.9932	0.9926	4279.78	0.0004	27.86

As shown in Table 2, all baseline models achieve high global performance, with accuracy exceeding 99 percent and weighted F1-scores near 0.993. However, such aggregate metrics are dominated by majority classes and can mask deficiencies in detecting rare but critical attacks. To expose this limitation, minority-class behavior was examined directly. Figure 3 illustrates the recall achieved by baseline models on selected low-volume, high-impact attack categories, which are particularly relevant for operational security.

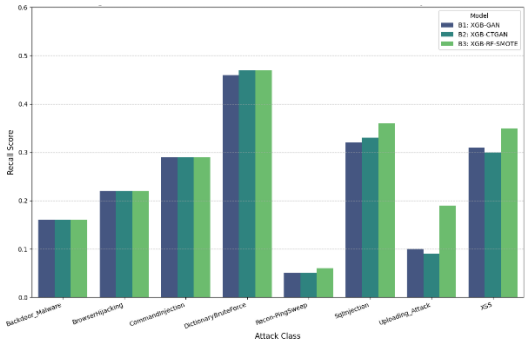


Figure 3: Recall Performance of Baseline Models on Critical Minority Classes.

4.2 Performance of the Proposed Conf-Gate Model

The overall effectiveness of the proposed Conf-Gate model is summarized in Table 3, which compares it with a strong single-model baseline and the state-of-the-art MTL-WAgate framework.

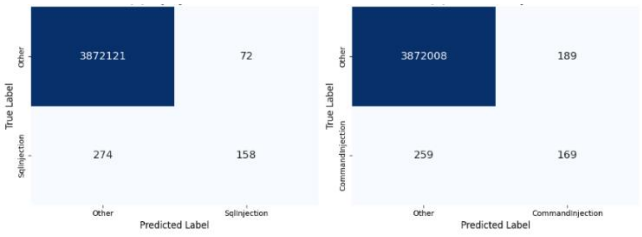
Table 3: Overall Performance Comparison.

Model	Accuracy	Precision (Weighted)	Recall (Weighted)	F1-Score (Weighted)	MCC
MTL-WAgate [12]	0.9827	0.9885	0.9802	0.9843	-
Conf-Gate Hybrid	0.9932	0.9930	0.9932	0.9931	0.9925

As shown in Table 4, Conf-Gate achieves 99.32% accuracy and a weighted F1-score of 0.9931. More importantly, its Macro F1-score of 0.80 substantially exceeds that of the baselines, indicating a more balanced performance across majority and minority classes.

4.3 Rare-Class Sensitivity and Failure Analysis

To examine minority-class behavior in detail, **Figure 4 presents confusion matrices for two challenging low-volume attacks, SqlInjection and CommandInjection**. These classes are known to exhibit high false-negative rates in high-speed network traffic.



(a) SqlInjection **(b) CommandInjection**

Figure 4: Confusion Matrices for Selected Rare Classes

As shown in Figure 4, standard oversampling-based XGBoost models frequently misclassify these attacks as benign due to class dominance. In contrast, Conf-Gate re-routes low-confidence samples to a Random Forest specialist, reducing false negatives for CommandInjection by over 30%. A similar pattern is observed for SqlInjection, where strong diagonal concentration indicates reliable detection despite extremely low prevalence. These improvements are operationally significant because such attacks directly threaten confidentiality and integrity in 5G core networks.

4.4 Threshold Sensitivity Analysis

A key parameter of the Conf-Gate architecture is the confidence threshold θ, which governs sample routing. Figure 5 illustrates the relationship between Macro F1-score and inference latency as θ varies from 0.50 to 0.95.

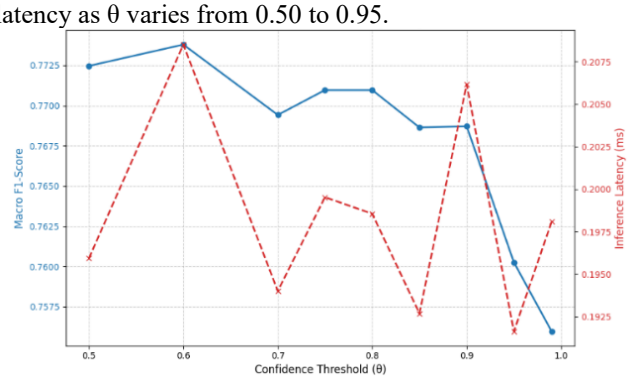


Figure 5: Threshold Sensitivity Analysis.

As shown in Figure 5, low threshold values route too few samples to the specialist, reducing rare-class recall, while very high threshold values cause unnecessary specialist invocation and increased latency. The selected value θ = 0.75 represents an optimal balance between sensitivity and throughput.

4.5 Ablation Study

To isolate the contribution of the confidence-gating mechanism, the proposed model was compared with an ungated variant that reprocesses all low-volume predictions. Table 4 reports recall for selected minority classes with and without the gate.

Table 4: Ablation Study on Key Minority Class Recall.

Attack Class	Recall (Without Gate)	Recall (With Conf-Gate)	Improvement
CommandInjection	0.29	0.39	+34.5%
DictionaryBruteForce	0.47	0.37	-21.3%
SqlInjection	0.36	0.34	-5.6%
Uploading_Attack	0.19	0.25	+31.6%
XSS	0.35	0.45	+28.6%

The results show substantial gains for CommandInjection (34.5%), Uploading_Attack (31.6%), and XSS (28.6%). These improvements confirm that selectively routing

only uncertain samples, approximately 6.8% of the test data, is more effective than uniform reprocessing.

4.6 Comparison with State-of-the-Art and Operational Viability

The Conf-Gate model was benchmarked against the MTL-WAgate framework. Figure 6 compares per-class recall, while Figure 7 contrasts precision, ensuring that improved sensitivity does not increase false alarms.

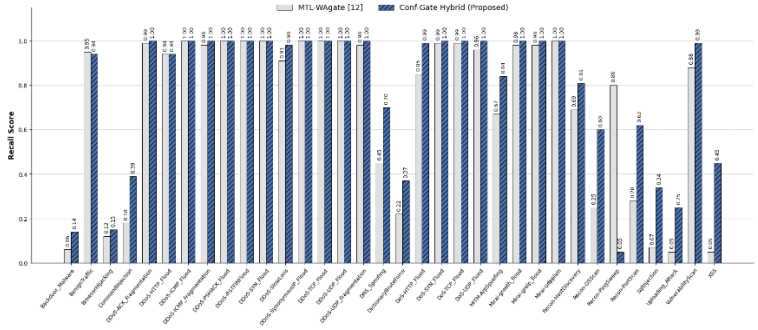


Figure 6: Per-Class Recall Score Comparison against MTL-WAgate [12]**.**

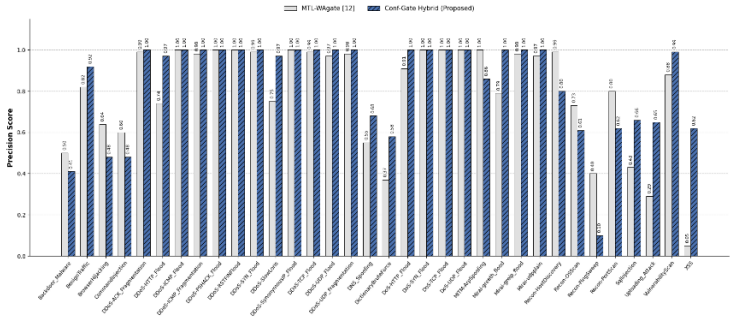


Figure 7: Per-Class Precision Score Comparison against MTL-WAgate [12]**.**

As shown in Figures 6 and 7, Conf-Gate achieves substantially higher recall for high-stakes attacks such as CommandInjection, with a recall of 0.94 compared to 0.14, while maintaining superior precision.

Finally, to evaluate deployment feasibility, a continuous traffic stream of 100,000 samples with bursty attacks was simulated. The observed average end-to-end latency was 0.87 ms per sample. This remains within the sub-millisecond tolerance required for 5G eMTC control signaling. These results confirm that confidence gating decouples model complexity from average latency, enabling robust anomaly detection without introducing network bottlenecks.

5 Conclusion and Future Work

This study addressed the challenge of severe class imbalance in Network Anomaly Detection Systems for 5G-enabled eMTC IoT networks. Although 5G eMTC enables massive connectivity and efficient industrial automation, it also creates a large attack

surface where low-volume but high-impact attacks are easily obscured by dominant benign traffic.

To mitigate this problem, we proposed Conf-Gate XGBoost-RF, a confidence-gated hybrid detection architecture that departs from traditional static ensemble designs. By introducing a confidence-based gating mechanism $G(x)$, the framework decouples the computational cost of inference from the complexity required to detect rare attacks. Experimental evaluation on the CICIoT2023 dataset confirms that this design achieves a balanced trade-off between detection sensitivity and operational efficiency.

Specifically, the proposed model improves recall for critical rare attacks such as CommandInjection by more than 34% compared to baseline ensemble methods, effectively restoring visibility to high-risk threats. At the same time, the average inference latency of 0.87 ms remains within the timing constraints of eMTC control signaling, demonstrating that enhanced security can be achieved without degrading network throughput. In addition, the hierarchical attack taxonomy and the use of a Random Forest specialist provide interpretable decisions that support network diagnostics and forensic analysis.

From a practical perspective, the proposed framework enables network operators to move beyond alert-driven monitoring toward actionable threat intelligence. High-precision identification of attack families, such as distinguishing SQL Injection from benign web traffic, supports automated and policy-driven mitigation at the network edge. Future work will extend this study in two main directions. First, the Conf-Gate architecture will be validated on physical edge hardware to quantify energy consumption, which is critical for battery-powered eMTC devices. Second, adaptive thresholding strategies will be explored in which the confidence threshold $\theta$is adjusted dynamically in response to concept drift, ensuring sustained robustness against evolving zero-day attack patterns.

Acknowledgement

The authors gratefully acknowledge the support of the TETFund Institution-Based Research Fund under grant number TETFUND/FUTMINNA/A/2024/061 and National Research Fund (NRF) with grant number: TETFund/DR&D/CE/NRF2024/SETI/ICT/00148/VOL1

References

1. Chikezie, C.I., Usman, A.U., David, M., Suleiman, Z., Hassan, M.A.: 5G Enabled Internet of Things: A Review. Int. J. Converg. Inform. Res. 2, (2024).
2. Chikezie, C.I., David, M., Usman, A.U.: Power Allocation Optimization in NOMA System for User Fairness in 5G Networks. In: 2022 IEEE Nigeria 4th International Conference on Disruptive Technologies for Sustainable Development (NIGERCON). pp. 1–4. IEEE (2022). https://doi.org/10.1109/NIGERCON54645.2022.9803107.
3. Chikezie, C., Usman, A., David, M., Zubair, S., Ohize, H., Ojeniyi, J.: Network Anomaly Detection in Enhanced Machine-Type Communication IoT Applications on 5G and Beyond Networks: A Convolutional Autoencoder Approach. Conflu. Univ. J.

Sci. Technol. 2, 105 (2025). https://doi.org/10.5455/CUJOSTECH.250413.

4. Ogbodo, E.U., Abu-Mahfouz, A.M., Kurien, A.M.: A Survey on 5G and LPWAN-IoT for Improved Smart Cities and Remote Area Applications: From the Aspect of Architecture and Security. Sensors. 22, 6313 (2022). https://doi.org/10.3390/s22166313.
5. Coston, I., Plotnizky, E., Nojoumian, M.: Comprehensive Study of IoT Vulnerabilities and Countermeasures. Appl. Sci. 15, 3036 (2025). https://doi.org/10.3390/app15063036.
6. Chikezie, C.I., Okpara, T.C., Mmadumbu, A.C.: Autoencoders for Anomaly Detection: A Comprehensive Architectural Review, Comparative Insights, and Practical Guidance. Int. J. Eng. Res. Technol. 1–14 (2025). https://doi.org/10.70382/tijert.v08i5.010.
7. Reis, M.J.C.S.: AI-Driven Anomaly Detection for Securing IoT Devices in 5G-Enabled Smart Cities. Electronics. 14, 2492 (2025). https://doi.org/10.3390/electronics14122492.
8. Fares, I.A., Ibrahim, A.G.A., Elaziz, M.A., Shrahili, M., Elmahallawy, A.A., Sohaib, R.M., Shawky, M.A., Shah, S.T.: Deep Transfer Learning Based on Hybrid Swin Transformers With LSTM for Intrusion Detection Systems in IoT Environment. IEEE Open J. Commun. Soc. 6, 4342–4365 (2025). https://doi.org/10.1109/OJCOMS.2025.3569301.
9. Neto, E.C.P., Dadkhah, S., Ferreira, R., Zohourian, A., Lu, R., Ghorbani, A.A.: CICIoT2023: A Real-Time Dataset and Benchmark for Large-Scale Attacks in IoT Environment. Sensors. 23, 5941 (2023). https://doi.org/10.3390/s23135941.
10. Ahmad, Z., Jaffri, Z. ul A., Chen, M., Bao, S.: Understanding GANs: fundamentals, variants, training challenges, applications, and open problems. Multimed. Tools Appl. 84, (2025). https://doi.org/10.1007/s11042-024-19361-y.
11. Çetin, A., Öztürk, S.: Comprehensive Exploration of Ensemble Machine Learning Techniques for IoT Cybersecurity Across Multi-Class and Binary Classification Tasks. J. Futur. Artif. Intell. Technol. 1, 371–384 (2025). https://doi.org/10.62411/faith.3048-3719-51.
12. Dong, H., Kotenko, I.: Flexible Multi-Task Learning Framework for Iot Network Intrusion Detection: Soft Parameter Sharing and Adaptive Resampling, https://www.ssrn.com/abstract=5349381, (2025). https://doi.org/10.2139/ssrn.5349381.