Future Technology
Open Access Journal
ISSN 2832-0379

August 2026| Volume 05 | Issue 03 | Pages 192-205

Journal homepage: https://fupubco.com/futech

https://doi.org/10.55670/fpll.futech.5.3.17

Future Publishing LLC

Article

Imbalance-resistant multiclass attack classification in real-time IoT water networks using SMOTE-enhanced random forests

Anita Anand[1]*, Shivangi Surati[2]

[1]Department of Computer Engineering, LDRP-ITR, Kadi Sarva Vishwavidyalaya, SVKM, Gujarat, India
[2]Department of Computer Science and Engineering, Pandit Deendayal Energy University, Gujarat, India

ARTICLE INFO

Article history:
Received 23 December 2025
Received in revised form
02 May 2026
Accepted 04 June 2026

Keywords:
Multiclass classification, Class imbalance, SMOTE,
Intrusion Detection System (IDS),
Smart water networks,
Water infrastructure resilience

*Corresponding author
Email address:
shivangi.surati@gmail.com

DOI: 10.55670/fpll.futech.5.3.17

ABSTRACT

The implementation of smart water distribution systems that rely on the Internet of Things (IoT) has substantially increased the need for intrusion detection systems capable of distinguishing among various categories of attackers. Such granularity is essential for timely and appropriate incident response. The nature of telemetry streams in operational settings is imbalanced: normal traffic is prevalent, whereas the rare but important classes of attacks are represented by a small number of attacks. In such circumstances, the traditional type of classifier can achieve high overall accuracy but fails to identify minority threats of greatest operational interest. This paper introduces a multi-class attack classification model that is robust to class imbalance and operates in real time on the IoT water network, classifying samples using the Synthetic Minority Over-sampling Technique (SMOTE) combined with a Random Forest (RF) ensemble classifier. The data used in the study is a collection of 1,048,575 telemetry records that simulate smart water infrastructure behavior by combining network indicators such as AnomalyScore, DataRate, and Protocol with physical-process indicators such as WaterFlowRate (Lpm), thereby covering cyber-physical interactions. An RF model trained on the original imbalanced dataset is compared with one trained on SMOTE-balanced data and evaluated on an unseen imbalanced test set. Even though the baseline achieves 99.3% accuracy, its recall is 0% for the rare DoS and DDoS classes. However, in comparison, the SMOTE-enhanced model obtains 99.88% accuracy and a higher recall of 92.31% for DoS and 99.66% for DDoS, and the macro- averaged F1-score rises from 0.60 to 0.93. The most discriminative features are recognized as AnomalyScore, DataRate, and WaterFlowRate (Lpm), which support interpretability and informed decision-making in sustainability-sensitive smart water infrastructure.

1. Introduction

IoT- based sensing systems are becoming more common in contemporary water distribution systems, where they continuously transmit telemetry data to facilitate leakage detection, demand balancing, and real-time operational monitoring [1,2]. Real-world implementations of smart water metering systems have shown that real-time feedback can alter consumer usage patterns, significantly increasing data volume and operational complexity [3]. As cyber-physical systems grow larger and more interdependent, the attacks they may be exposed to become more frequent, and the need for reliable operation to safeguard vital services increases [4,5]. Machine learning (ML) has become a prevalent means of detecting threats in IoT settings and offers greater flexibility than rigid signature-based protection against evolving attack patterns [6]. Anomaly detection surveys on smart water systems have emphasized the application of Random Forests, Support Vector Machines, k-Nearest Neighbors, and autoencoder-based algorithms, which have mostly been evaluated using metrics such as precision, recall, F1-score, and accuracy [7]. Recent research on smart-city water infrastructure also highlights the importance of implementing a robust detection system to reduce cyber incidents and operational failures [8]. However, in binary benign-malicious classification is not enough in operational security environments. The response

in a Security Operations Center (SOC) necessitates identifying multi-class attacks, enabling analysts to recognize Distributed Denial-of-Service (DDoS), Denial-of-Service (DoS), botnet activity, and injection-based manipulation, rather than relying on general anomaly notifications [9]. Achieving this level of detection granularity is challenging due to a defining characteristic of real-world network telemetry: severe class imbalance, in which benign traffic predominates while certain attack categories occur only rarely. Under such conditions, machine learning models may optimize overall accuracy yet exhibit poor recall for minority classes, including the most critical security threats [10]. To address this challenge, the study proposes an imbalance-resistant multi-class attack classification framework for real-time IoT water telemetry by integrating SMOTE with a Random Forest ensemble classifier. SMOTE synthesizes minority-class samples in the feature space, enabling the classifier to learn reliable decision boundaries for rare attack categories [11]. Random Forest ensembles are widely adopted in IoT environments for intrusion detection frameworks owing to their ability to handle model complexity and heterogeneous tabular data distributions, while simultaneously offering intrinsic feature-importance metrics that enhance interpretability and support decision-making [12]. Unlike other previous investigations, which are largely concerned with aggregate performance improvements, the article is an experimental comparison between a Random Forest (RF) model trained on imbalanced data and the same model trained on SMOTE-balanced data, which allows us to specifically quantify the impact of class balancing on minority class recall and macro-averaged performance measures.

2. Related work

Machine learning techniques have become fundamental to strengthening the security posture of networked and cyber-physical systems, particularly within IoT deployments. As IoT environments evolve from localized smart home installations to large-scale, mission-critical infrastructures, operators increasingly rely on intelligent Intrusion Detection Systems (IDS) capable of identifying diverse attack patterns in real time. A substantial body of research demonstrates that ML-based detection approaches outperform traditional signature-driven methods when identifying malicious or anomalous activity embedded within complex traffic streams [11].

2.1 Ensemble methods for IoT intrusion detection

Enclosed in the extended machine learning environment, it has become a solid methodology for developing high-performance IDS solutions. Ensemble models combine predictions of several base learners, ensuring higher predictive accuracy, stronger generalization to unseen data, and reduced vulnerability to overfitting on noisy training data. Many methods have been investigated across various IoT settings, including military communication networks and heterogeneous civilian deployments. Bagging, boosting, and stacking techniques have been discussed in various settings [12-14]. Random Forest (RF) is a bagging-based ensemble of decision trees that is widely used for its strong empirical performance and relatively interpretable behavior compared to other methods [15,16]. Prior research has shown that RF, when combined with effective feature engineering and selection techniques, is highly effective at detecting advanced threats such as botnets and distributed denial-of-service (DDoS) attacks [17,18]. Ensemble models can better handle the high dimensionality, noise, and heterogeneity of IoT traffic than most single-model options [19].

2.2 Handling severe class imbalance

Despite their effectiveness, machine learning models frequently struggle with the severe class imbalance characteristic of cybersecurity datasets. In real network traffic, benign activity constitutes the overwhelming majority, while attack samples are sparse and unevenly distributed across threat categories. As a consequence, models may achieve deceptively high accuracy by predominantly predicting benign labels while failing to detect rare yet high-impact intrusions [20]. This imbalance-induced blind spot poses a significant barrier to deploying trustworthy IDS solutions in operational environments [21]. To address this challenge, prior research has explored both data-level and algorithm-level remedies. Among data-centric approaches, the Synthetic Minority Over-sampling Technique (SMOTE) and its variants are widely recognized as effective tools for balancing skewed datasets [22,23]. Rather than duplicating minority samples, SMOTE generates synthetic instances by interpolating between neighboring minority points in feature space, enriching minority distributions and enabling clearer decision boundaries. Numerous IoT security studies validate the effectiveness of SMOTE for intrusion detection [24], including specialized research on botnet traffic and hybrid frameworks that combine resampling with additional optimizations [25]. Domain-specific extensions further refine synthetic sample generation for resource-constrained IoT edge environments [26]. Collectively, recent literature emphasizes that explicit imbalance mitigation is essential for IDS designed to detect infrequent but high-impact attacks [27].

2.3 Domain challenges in smart water networks

Cyber-physical systems involving smart water metering and broader water distribution pose domain-specific challenges that complicate intrusion detection. In addition to traditional cyberattacks, operators must also contend with sensor failures, hydraulic anomalies, and intentional actions targeting valves, pumps, or meter readings. Machine learning-based anomaly detection is increasingly used in water CPS environments to identify deviations from normal hydraulic behavior. Recent research highlights the need for scalable and accurate detection frameworks capable of processing both benchmark datasets and real-time telemetry streams from smart meters and supervisory control systems [28]. In practice, IDS designs should achieve high detection performance under real-world conditions, including limited bandwidth, measurement noise, and non-stationary demand patterns. Furthermore, they must effectively address class imbalance to ensure that rare but safety-critical attacks are reliably detected rather than being overshadowed by the overwhelming volume of benign traffic.

2.4 Research gap and contribution

According to the reviewed literature, research on intrusion detection for smart water meters has primarily focused on employing various machine learning and deep

learning algorithms. The reviewed literature shows that SMOTE and ensemble learning are popular approaches for intrusion detection systems; however, current research lacks a controlled experimental design to isolate the specific contribution of data rebalancing to minority-class detection. Most approaches embed SMOTE within complex pipelines, making it difficult to attribute performance improvements to resampling alone. Furthermore, most research has ignored potential threats to smart water meter infrastructure, such as DoS and DDoS attacks, in favor of identifying anomalies within the digital domain, such as network traffic or sensor data. This study attempts to fill these gaps by performing a controlled comparison to explicitly quantify the effect of SMOTE on the recall of the minority class and macro-level performance. This article presents an empirical comparison of a baseline random forest trained on imbalanced data and the same model trained on a SMOTE-balanced dataset. Particular emphasis is placed on quantifying improvements in the recall of rare attack categories and assessing the impact of class balancing on overall detection performance. Table 1 situates the pipeline within the prior literature. Although previous works use advanced ensembles and improved resampling strategies, our framework is unique in that it provides a validated need for SMOTE through a controlled baseline rather than enhanced analysis. The findings indicate that class balancing converts the completely undetected DDoS and DoS classes into reliably detected ones, with recall exceeding 92%, hence proving that systematic data balancing is a significant component of contemporary, reliable IDS in intelligent water infrastructures.

3. Problem definition

3.1 System and threat model

In the article, the smart water metering network is discussed as an example of a cyber-physical system in which IoT-based sensors continuously measure water flow, pressure, and consumption patterns. Such deployments produce telemetry streams at high velocity, including both network-layer indicators (i.e., DataRate and Protocol) and physical-domain metrics (i.e., WaterFlowRateLpm). Because digital control infrastructure and physical processes are highly interdependent, they are also vulnerable to cyberattacks that might disrupt operations, compromise sensor integrity, or obstruct authorized access to data**.** The proposed approach introduces an intrusion detection system design to identify four main categories of threats in a smart water meter environment. (1) Denial-of-service (DoS) where attackers purposefully overload the target system by flooding it with traffic, making it inaccessible to legal users. It primarily targets IoT sensor devices, leading to data disruption and meters stopping the sending of readings. (2) Distributed-Denial-of-Service (DDoS) is an advanced version of a DoS attack that exhausts the IoT network resource at the network layer, leading to system failure and making it impossible to monitor water distribution across the city [29]. (3) SQL Injection is detected at the data layer, which compromises the security, data integrity, and operational dependability of smart city systems. Leads to manipulation of billing records [30]. (4) Web Attacks target the application layer, which allows attackers to use application dashboards and results in weak authentication, session hijacking, disruption of monitoring services, and exposure of sensitive customer information [31].

3.2 Class imbalance in multiclass

Consider a dataset defined as $D = \{(x_i, y_i)\}^N$, where each $x_i \in R^d$ represents observed network features and $y_i \in \{c_1, c_2, \ldots, c_k\}$ denotes the corresponding threat class. Operational IoT traffic exhibits pronounced class imbalance: benign observations N_{maj} vastly outnumber attack samples $N_{min,j}$ across minority classes,

$$N_{maj} \gg N_{min,j}, \quad \forall j \in \{1, \ldots, k-1\}$$

Table 2 shows the empirical class distribution of the smart water meter dataset, which clearly depicts the severity of class imbalance. While benign traffic accounts for more than 51% of the dataset, critical attack classes such as DDoS and DoS constitute only 0.62% and 0.01%, respectively. The imbalance ratio (IR) is calculated as the ratio of the majority class size to the size of the specific class IR = size(majority class)/size(class). An IR of 1 indicates the majority class, while higher values indicate a severe minority class. The IR between benign and DoS exceeds 3433:1, confirming the condition $N_{maj} \gg N_{min,j}$. This perfectly highlights the necessity of using oversampling techniques like SMOTE to prevent the model from ignoring this severe minority class.

3.3 Research objective

The primary objective of the article is to develop and evaluate a machine learning framework to detect minority intrusion classes with high accuracy in IoT-based smart water distribution systems, in the presence of severe class imbalance, with a particular focus on improving minority-class recall. The secondary goal is to achieve balanced classification performance across all attack categories while maintaining high overall accuracy.

4. Methodology

This section presents the entire pipeline for constructing a multiclass attack model that remains resilient to extreme class imbalance in IoT-based smart water systems. The proposed framework combines distributed data processing, heuristic label generation, targeted class rebalancing with SMOTE, and classification with an ensemble of classifiers via a random forest. Evaluation of the pipeline is performed through a controlled comparison between the baseline and enhanced settings.

4.1 System architecture

The proposed framework consists of a hybrid data preprocessing architecture that combines Apache Spark and SciKit-learn to balance scalability and model flexibility. The dataset includes large-scale water telemetry (1,048,575 records) processed in a batch-oriented manner using PySpark for offline training and evaluation. The total PySpark execution time was 26.96 seconds for 1,048,575 records, with an ingestion rate of ~38,893 events per second. In contrast to Apache Spark, which handles large-scale distributed processing of in-memory data, classical ML libraries like SciKit-learn and imbalanced-learn work with in-memory data structures. Thus, when distributed feature engineering is performed, the dataset is converted to a Pandas DataFrame using the .toPandas() method to effectively train a model and apply SMOTE.

Table 1. Comparison with existing methods

Approach	Methodology	Focus Area	Class Imbalance Handling	Comparative Analysis (Baseline vs. Enhanced)	Key Contribution	Overall Accuracy
Thakkar & Lohiya [10]	Ensemble Deep Learning	Attack Classification in IoT	Acknowledged (addressed with ensemble)	No (Presents final model performance)	A high performance deep learning framework for imbalanced IoT data	98.5%
Manokaran et al. [26]	Fuzzy Clustering + SMOTE	Anomaly Detection at IoT Edge	Yes (Proposed novel "PPFCM-SMOTE")	No (Focuses on novel SMOTE variant's performance)	A new specialized SMOTE technique for IoT edge environments	97.8%
Saranya&Priyad- Harshini [3]	KD-SMOTE + Ensemble	CyberThreat Detection	Yes (Proposed novel "VAST-KD-SMOTE")	No(Compared against other optimizers, not non-SMOTE baseline)	A dual-strategy framework optimizing balancing and feature selection	99.1%
Bokhari et Al. [8]	Bagging Ensemble + SMOTE	General IoT Cyberattack Detection	Yes (SMOTE applied)	No(Presents final performance of enhanced model)	Demonstrates effectiveness of Bagging ensemble on SMOTE-balanced dataset	99.4%
Our Proposed Work	Random Forest + SMOTE	High-Recall MultiClass Classification in Smart Water Networks	Yes (Standard SMOTE)	Yes (Direct Quantitative comparison of with-SMOTE vs.without SMOTE)	Empirically proves SMOTE necessity - recall lift from 0% to >92%	99.88%

Table 2. Attack distribution count with imbalance ratio

Attack Type	Count	Percentage	Imbalance Ratio (IR)
Benign	535,689	51.09 %	1.00
Web Attack	327,587	31.24 %	1.64
SQL Injection	178,606	17.03 %	3.00
DDoS	6,537	0.62 %	81.95
DoS	156	0.01 %	3,433.90
Total	1,048,575	100.0%	

This step collects deeply compressed, preprocessed data from the distributed cluster into a single optimized local Pandas DataFrame. During the first phase of data ingestion, the RecordID is kept to enable forensic analysis and system traceability. It is excluded from the model's feature vectors during training to avoid bias, but kept separately and linked back to the prediction outputs. This allows each prediction to be traced back to its original data record, enabling detailed investigation of events such as anomalies or attacks. The dataset uses stratified random splitting (70:30) for evaluation, while RecordID is preserved and reattached to predictions to ensure full traceability. The whole imbalance-resistant intrusion detection pipeline is represented in Figure 1. It demonstrates the flow from telemetry acquisition and SMOTE-based model training to real-time deployment. The pipeline, supported by a hybrid analytics stack comprising distributed preprocessing and in-memory machine learning, is shown in Figure 2.

Figure 3 presents the detailed workflow. Preprocessing and heuristic label synthesis are followed by stratified train–test partitioning to ensure unbiased evaluation. This method deliberately used a 70:30 stratified split to ensure that the extreme minority-class DoS attacks, representing 0.01% of the data, were proportionally represented in both the training and testing phases, prioritizing multi-class detection accuracy over temporal sequence modeling. Two parallel training paths are then executed: a baseline Random Forest trained on the original imbalanced data and a proposed Random Forest trained on SMOTE-rebalanced data. Both models are evaluated on a test dataset to ensure a fair comparison.

4.2 Dataset preparation and feature handling

The dataset consists of 1,048,575 records generated by Internet of Things smart water meter infrastructure from the urban area of Ahmedabad city, Gujarat, India, with 14 original features spanning a continuous operational period of around 4 weeks in the year 2022, in the month of August. The dataset used in this research is proprietary and provided by the Ahmedabad Municipal Corporation, and, due to operational and privacy constraints, it is not publicly available. The real-time dataset consists of data collected in West Ahmedabad by network sniffers and physical sensors at edge IoT gateways in every household. The dataset reflects the operational nature of IoT-based cyber-physical monitoring systems. Table 3 presents the features of the smart water meter dataset for households in West Ahmedabad.

(1) Feature Selection: Feature selection was performed using the Random Forest-based importance ranking approach, where feature relevance was measured using the intrinsic Mean Decrease in Impurity (MDI), commonly referred to as "Gini Importance".

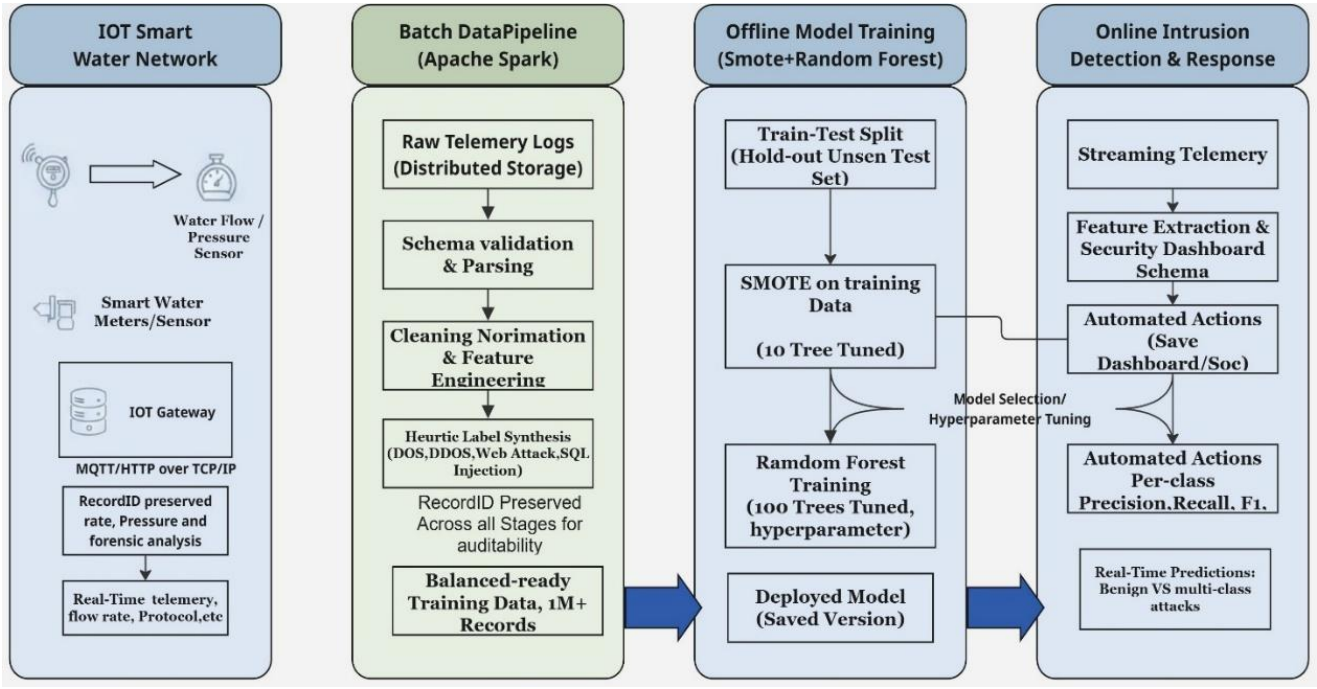

Figure1. Full architecture for imbalance-resistant intrusion detection in Internet of Things smart water networks. Flow and pressure sensors provide real-time telemetry that is relayed through an IoT gateway. Spark is used to verify, clean, and normalize data, perform feature engineering, and heuristically label and assign persistently unique RecordIDs to make things traceable. In offline training, SMOTE is applied only to the training split, and then a Random Forest model is fitted on an imbalanced test set. The chosen model is implemented online to use real-time multiclass predictions, and mitigation is implemented automatically.

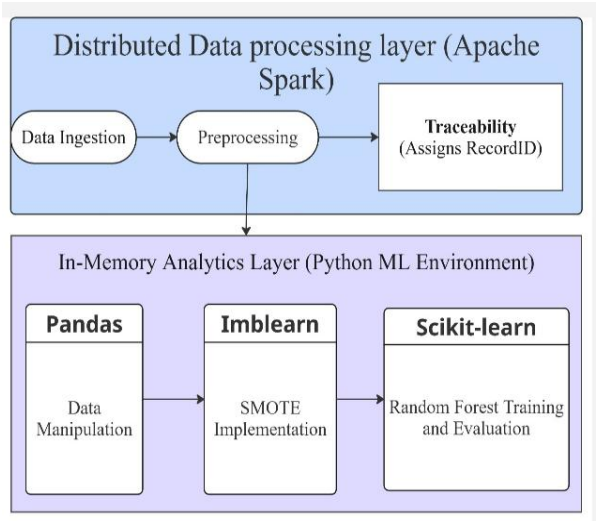

Figure 2. Hybrid data analytics architecture combining distributed preprocessing with in-memory model training and evaluation

The top nine features selected by the Gini Importance algorithm after applying SMOTE balancing are explicitly ranked as shown in Figure 4. These features constitute the optimal set of cybersecurity telemetry requirements for multiclass intrusion detection.

(2) Heuristic Label Assignment: Since operational environments do not provide expert-labeled datasets of cybersecurity, multiclass ground truth labels were created based on a deterministic rule-based framework, which is domain knowledge-based, as shown in the following algorithm.

Algorithm: Deterministic Multi class label assignment

Input: Telemetry row (R) containing cyber-physical and network metrics.

Output: Ground Truth label (L) ∈ {DoS,DDoS,SQL Injection,Web Attack}

If R.AttackPrediction==1 AND R.AnomalyScore> 0.7 then
L ← "DDoS"
Else If R.WaterFlowRate_Lmp < 3.0 then
L ← "SQL Injection"
Else If R.DataRate > 30.0 then
L ← "Web Attack"
Else If R.AttackPrediction==1 then
L ←"DoS"
Else
L ← "Benign"
End If
Return L

AnomalyScore is a continuous outlier metric computed using the Isolation Forest classifier, which isolates outliers in the physical metrics with contamination=0.05. It computes the AnomalyScore using iso.score_samples() and then normalizes the output between 0 (normal) and 1 (anomalous) using min-max scaling. AttackPrediction is trained using the ground-truth rule: if the AnomalyScore is greater than 0.7, it is considered an attack and flagged as such between 0 and 1 by a supervised Random Forest that analyzes network traffic.

The result of this labeling process was a realistic four-class attack dataset with a strong imbalance, in which benign traffic predominantly represented the rare attack classes.

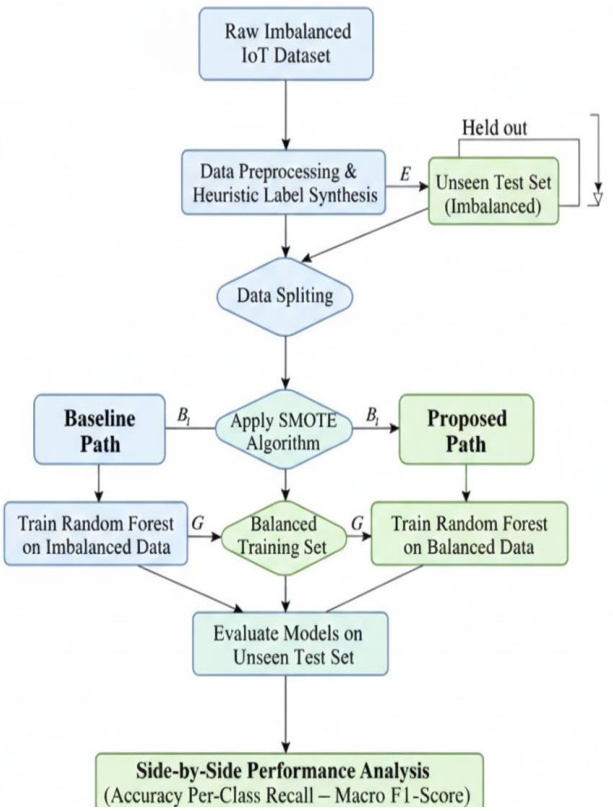

Figure 3. End-to-end methodology workflow from telemetry ingestion and heuristic labeling through SMOTE-based rebalancing, model training, and evaluation

Justification of Domain Attacks: DDoS – A highly generalized anomaly score, coupled with a positive attack prediction, leads to a massive, distributed volumetric attack that overwhelms the network gateway. SQL Injection: an extreme drop in physical water flow (less than 3 liters per min) leads to database-level actuators being maliciously manipulated. Web Attack - A network data rate exceeding 30 units indicates excessive payload requests typical of targeted web application vulnerabilities. DoS – if a general attack is flagged but none of the extreme multivariable thresholds above are met, it sound llike a single source denial-of-service attack disrupting localized sensor communication.

(3) SMOTE-Based Rebalancing: In order to reduce the distorted proportion of classes, the Synthetic Minority Over-sampling Technique (SMOTE) [32] was used. SMOTE creates artificial minority examples by interpolating the closest examples in a feature space, increasing the size of sparse minority groups, and enhancing class-separable boundaries. To avoid information leakage and maintain evaluation realism, SMOTE was applied only to the training split, and the test set retained its naturally imbalanced distribution, representative of real-life traffic.

(4) Random Forest Classification: Random Forest was chosen as the main classifier because it is rather strong and does not overfit. It can also model nonlinear relationships in structured telemetry data. The ensemble builds many decision trees using bootstrap aggregation and randomized feature selection, enhancing generalization and stability [33]. To make the model reproducible, it was run with n_estimators=100 and random_state=42.

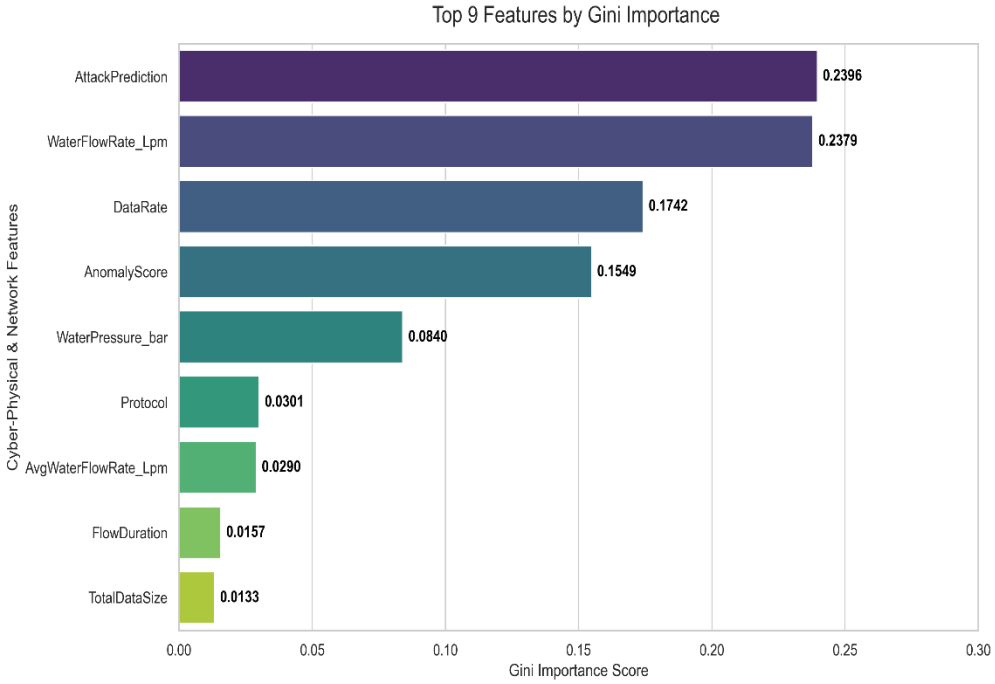

Figure 4. Feature Importance with MDI score

Table 3. Total features of the smart water meter dataset

Sr.No	Features
1	FlowDuration
2	Protocol
3	SessionDuration
4	DataRate
5	TotalDataSize
6	WaterFlowRate_Lpm
7	TotalWaterConsumed_L
8	AvgWaterFlowRate_Lpm
9	WaterUsageDuration_min
10	AmbientTemp_C
11	AmbientHumidity_percent
12	WaterPressure_bar
13	WeatherCondition
14	UsageFrequency_perDay

(5) Evaluation Protocol: The design of the study was a paired experiment to measure the effects of rebalancing the classes:
Baseline Model: Trained on the original unbalanced training dataset without any alteration to the distribution of classes using a Random Forest classifier.
Enhanced Model: Trained on a SMOTE-balanced selection of the training data implemented via the imbalanced-learn framework. Configured with a sampling_strategy='auto' and k_neighbors=5, where sampling_strategy generates synthetic records for all non-majority classes until uniform equilibrium is achieved, k_neighbors calculates Euclidean distances between feature vectors to ensure the generated samples rigorously adhere to the true topological boundaries of the minority attacks. Random Forest classifier, in which synthetic samples are created to better represent the minority classes [33].
A stratified train-test split of 70:30 preserves the natural class proportions in the test set, as is the case in practice. Performance analysis includes overall accuracy and per-class accuracy, recall, F1-score, and macro-averaged F1-score. Macro-F1 is preferable because it assigns equal importance to each class and provides a more accurate measure of performance in the presence of class imbalance.

5. Experiments and results

This section presents an empirical evaluation of class imbalance mitigation for multiclass intrusion detection in smart water IoT networks. The performance of a baseline Random Forest (RF) model trained on the original imbalanced dataset is compared with that of an identical RF model trained on SMOTE-rebalanced data. Confusion matrices, predicted class distributions, and classwise performance metrics are analyzed to determine whether resampling improves detection of rare threats without compromising performance on majority classes.

5.1 Experimental configuration

(1) Dataset Partitioning: A large smart water telemetry dataset (1,048,575 records) was used to conduct the experiment. A stratified 70:30 train-test split was used to maintain the natural class distribution in the test set. Configuration reflects realistic operational conditions, where benign traffic is the majority, and attacks are comparatively infrequent across operational conditions or environments.

(2) Implementation Details: All models were implemented in Python 3.9 with the Scikit-learn library to develop the models and the imblearn toolkit to use the SMOTE-based resampling. Tests were carried out on a 14-core CPU with 16 GB RAM .
(3) Model Variants and Comparison Design: To isolate the impact of data rebalancing, both model variants were configured with identical hyperparameters (n_estimators=100, random_state=42):
Baseline RF: The original imbalanced training data were used to train the baseline RF.
SMOTE+RF: trained on the SMOTE-balanced version of training data with the same training split.

5.2 Baseline RF (No SMOTE) Vs SMOTE Enhanced (SMOTE+RF) Performance Analysis

To avoid inaccuracies generated by skewness in the distribution of classes, the evaluation will focus on classwise precision, recall, F1-score and macro-averaged F1 (macro-F1). The macro-average assigns the same level of importance to each of the classes thus it becomes especially appropriate in the evaluation of performance in case of extreme class imbalance.
(1) Detailed Classification Metrics: Table 4 shows the comparison evaluation between baseline RF and SMOTE-enhanced RF models. The baseline RF model even though it is highly accurate (99.30%), achieves zero recall and F1-score for minority classes. This clearly demonstrates that accuracy alone is insufficient for evaluating intrusion detection systems under severe class imbalance.

Table 4. Classification report of baseline(no SMOTE) vs SMOTE+RF model

Attack Type	Model	Precision	Recall	F1-Score	Support
Benign	BaseLine	99.00	100.0	99.00	160707
	SMOTE+RF	99.99	99.86	99.93	535,689
Web Attack	BaseLine	99.00	99.00	99.00	98276
	SMOTE+RF	100.00	99.90	99.95	327,587
SQL Injection	BaseLine	100.00	99.00	99.00	53582
	SMOTE+RF	100.00	99.90	99.95	178,606
DDoS	BaseLine	0.00	0.00	0.00	1961
	SMOTE+RF	85.21	99.66	91.87	6,537
DoS	BaseLine	0.00	0.00	0.00	47
	SMOTE+RF	58.30	92.31	71.46	156
Macro Avg	BaseLine	0.60	0.59	0.60	314573
	SMOTE+RF	0.89	0.98	0.93	1,048,575
Accuracy	BaseLine	99.30%			1,048,575
	SMOTE + RF	99.88%			

In contrast, the SMOTE-enhanced RF model significantly improves minority-class detection while preserving near-perfect performance on majority classes. Recall improves from 0% to 99.66% for DDoS and from 0% to 92.31% for DoS. Similarly, macro-F1 improves from 0.60 to 0.93, indicating

substantially more balanced performance across all attack categories.

(2) Confusion Matrix Evidence: The effect of class imbalance on intrusion detection performance is shown in Figure 5 and Figure 6. The confusion matrix of the baseline RF shows no true positive detections for DDos and DoS, and predictions are biased towards the majority classes. On the other hand, SMOTE+RF exhibits high diagonal entries across all classes, reaffirming that balanced training data significantly improves the detection of rare attack types.

(3) Predicted Distribution Analysis: The predicted class distribution shown in Figure 7 further reflects majority bias. Most instances are classified as *Benign* or *Web Attack* with minimal predictions for *SQL Injection* and almost none for *DoS* or *DDoS*. This distribution confirms that the baseline model rarely identifies rare threat categories.

Figure 8 presents the predicted class distribution for the SMOTE-enhanced model. All attack categories receive substantial prediction counts, and *DDoS* and *DoS* predictions increase dramatically compared with the baseline, reflecting improved sensitivity to rare threats.

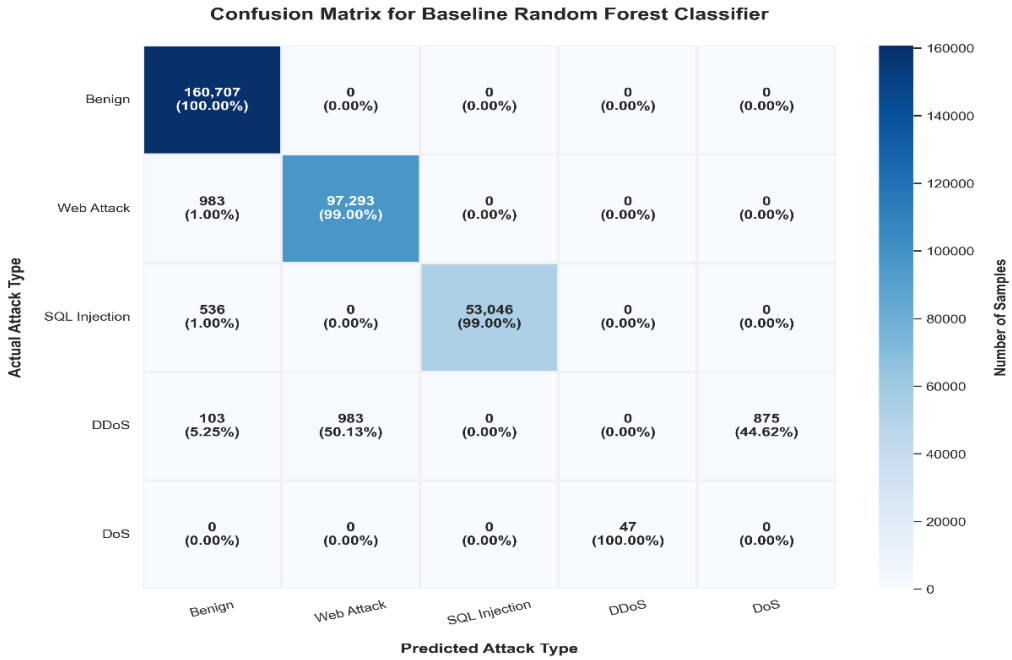

Figure 5. RF confusion matrix on the imbalanced test set

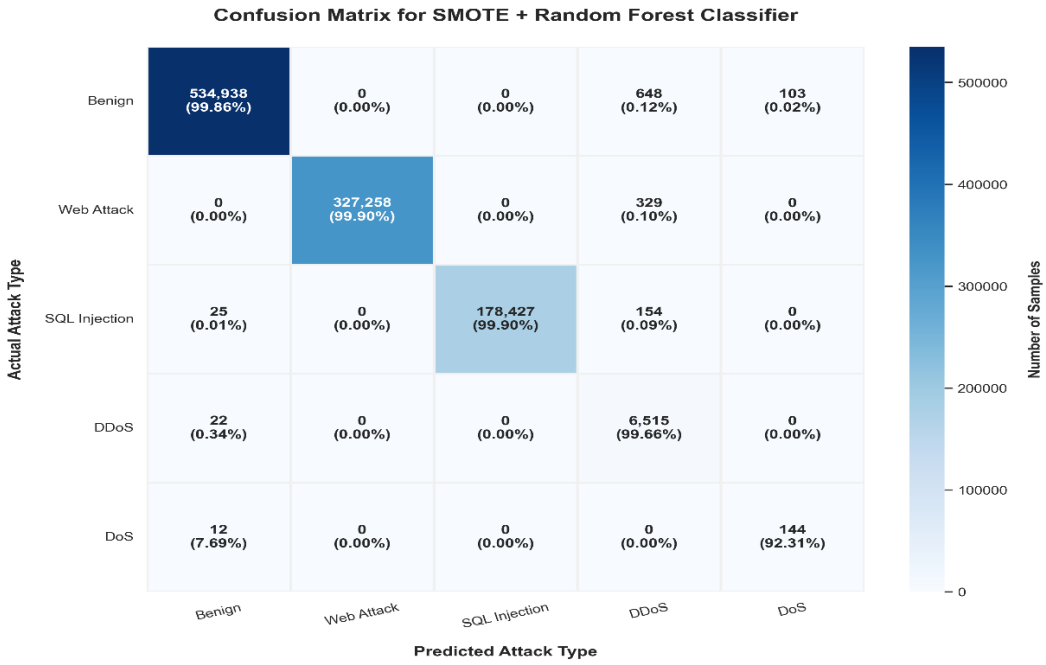

Figure 6. SMOTE+RF confusion matrix on the balanced test set. Rare attack types now exhibit strong true-positive counts

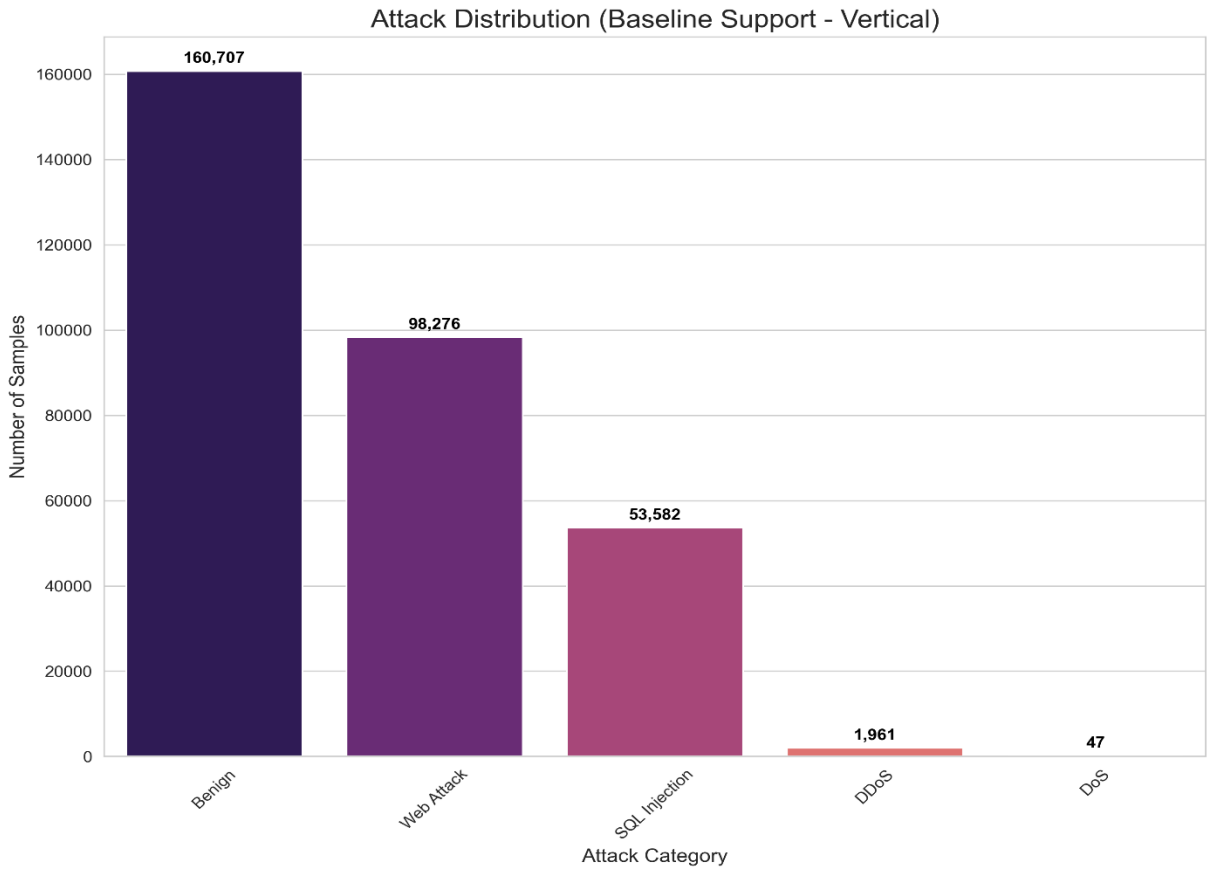

Figure 7. RF predicted attack distribution. Predictions are heavily skewed toward majority classes, with negligible DoS/DDoS output

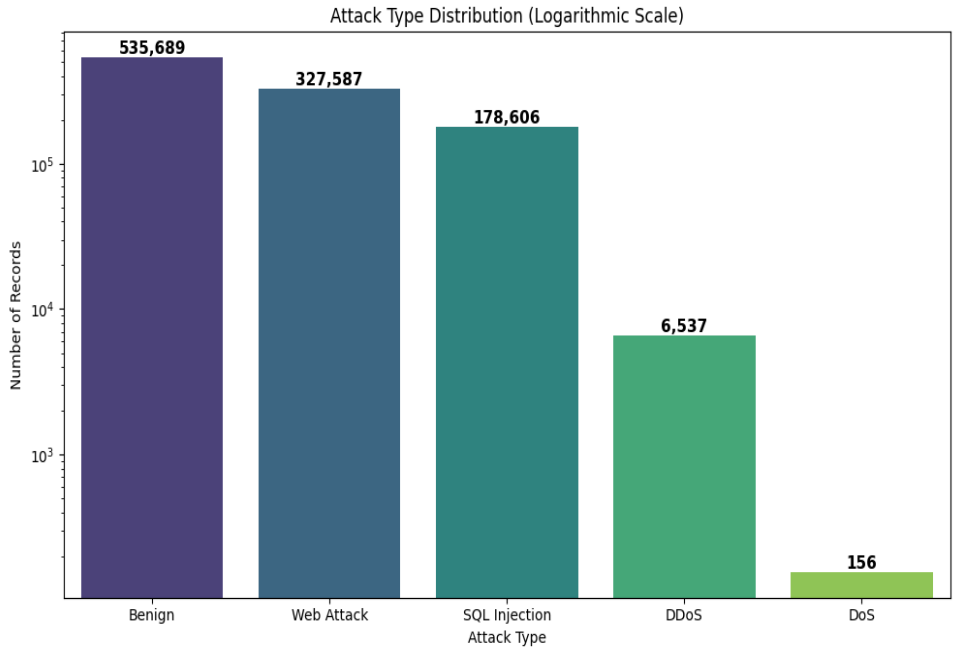

Figure 8. SMOTE+RF predicted attack distribution. All attack types, including rare DDoS and DoS, are regularly detected

5.3 Granular metric comparisons

(1) Recall: Threat Coverage: The recall comparison in Figure 9 highlights the contrast between the two models. Recall for *DDoS* and *DoS* collapses to zero in the baseline model, whereas the SMOTE-enhanced model maintains high recall across all classes, substantially reducing false negatives for rare threats.

(2) Precision: Alert Reliability: Figure 10 shows that precision for the majority classes remains near perfect in both models, while precision for the minority classes decreases modestly under SMOTE. This reflects a deliberate trade-off: increased sensitivity to rare events may introduce additional alerts but reduces the likelihood of missing critical incidents.

(3) F1-Score: Holistic View: Figure 11 shows the F1-score comparison, which is the product of precision and recall. The baseline model attains an F1 score of zero in both conditions of F1 of DDoS and F1 of DoS, but the SMOTE-enhanced F1 scores of the baseline model show good performance across all classes, improving the macro-F1 score from 0.60 to 0.93.

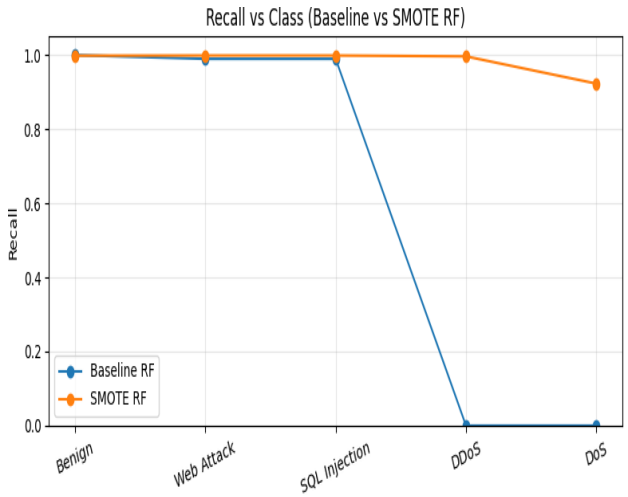

Figure 9. Per-class recall for Baseline RF and SMOTE+RF. SMOTE greatly enhances coverage of rare attack types

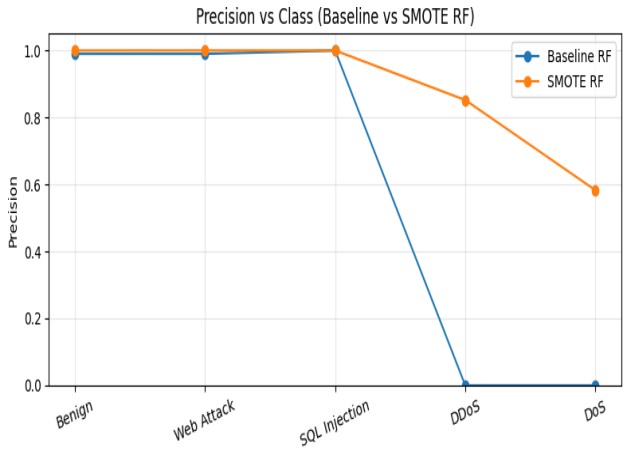

Figure 10. Per-class precision for Baseline RF and SMOTE+RF. The majority precision is preserved, while the minority precision reflects increased sensitivity

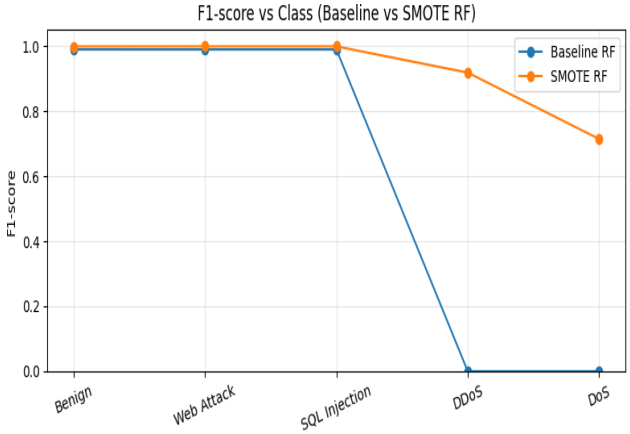

Figure 11. Per-class F1-score for Baseline RF and SMOTE+RF. SMOTE yields strong minority-class F1 while preserving majority performance

6. Discussion

The research findings shed light on the harmful effects of class imbalance on multiclass intrusion detection in IoT-based smart water networks. Aggregate accuracy is the sole indication that is unreliable for identifying rare but critical threats that might go undetected. However, SMOTE-based rebalancing reallocates this dynamism so that the detector can maintain the same high level of accuracy while significantly increasing recall across all attack types, including the least frequent ones.

6.1 Baseline limitations under skew

Despite the baseline Random Forest model having an overall accuracy of 99.3%, it has zero recall on both DoS and DDoS attacks, which are the most unusual and potentially the most harmful types of threat in the dataset. It is the result of a familiar effect of extreme class imbalance, namely that by emphasizing the overall error, the classifier optimizes (the error due to) the dominant benign class, at the expense of the few samples that are the minority and whose contribution to the overall error is marginal: that is, they carry an insignificant loss to the classifier. Such behavior makes the baseline not suitable to be deployed in terms of its operational security. The system that detects intrusions, which continuously ignores the high-impact attacks regardless of its claimed accuracy, cannot afford to protect the water infrastructure (critical) meaningfully.

6.2 Impact of SMOTE-based rebalancing

Applying SMOTE to rebalance the training data preserves near-perfect accuracy (99.88%) while substantially improving minority-class detection. DDoS recall goes up by 0%-99.66%, DoS recall goes up by 92.31%, and macro-F1 goes up by 0.60 to 0.93. These findings illustrate that information-level rebalancing is not only advantageous but also important for multiclass intrusion detection systems, which need to detect extremely rare attack types. Working on the principle of synthesizing minority-class samples by interpolating between neighboring samples in the feature space, SMOTE functions. The process of enriching the sparse areas around the rare attacks allows the Random Forest ensemble to learn more representative decision boundaries,

which can be used to predict the unseen data, in line with well-known resampling theory [22].

6.3 Security-relevant precision–recall trade-offs

The SMOTE-enhanced model demonstrated a favorable precision-recall trade-off, achieving substantially higher recall for minority attack classes while introducing a moderate increase in false positives. In critical infrastructure environments, the cost of missing a cyberattack is typically greater than the cost of additional alerts. Consequently, prioritizing minority-class detection is a practical and security-driven design choice for an intrusion detection system [11]. This trade-off aligns with real-world security priorities, in which overall threat reporting is usually favored over minimizing manageable false positives [20]. However, the added false positives from the proposed application of SMOTE for the majority classes cause a slight drop in precision from 99.00% to 99.99% for benign traffic and from 99.00% to 100% for WebAttack detection, while minority classes precision increase for DoS increases to 58.30% due to increased sensitivity toward rare attacks. This trade-off yields substantial gains in recall with DDOS improving from 0.00% to 99.66%, and DoS recall improving from 0.00% to 92.31%. In smart water meter infrastructure, the operational cost of missing a DoS or DDOS attack, such as SCADA communication failure, sensor outages, billing disruption, or denial of water services, is significantly higher than the cost of additional false-positive alerts [29,30].

To assess the performance of the classifier with severe class imbalance, two types of analysis were conducted:
ROC and precision-recall analysis. The imbalance introduced itself via bias, with the baseline RF model, as shown in Figure 12, being separable for majority classes but poorly discriminative for attacks on minority classes, especially DoS attacks. The results showed that minority class separability and recall significantly improved after applying SMOTE, thereby confirming improved learning of rare attack patterns, as shown in Figure 13.

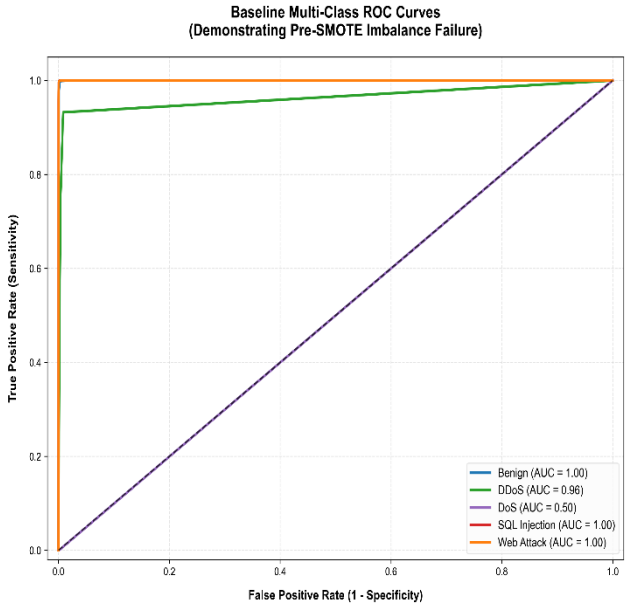

Figure 12. ROC curve for baseline model

As shown in Figure 14 and Figure 15, a precision-recall analysis also revealed that SMOTE boosted the accuracy of minority-class detection with a moderate increase in false positives. This trade-off is operationally acceptable in smart water intrusion detection systems, where minimizing missed attacks is more critical than reducing manageable false alarms. Overall, the ROC and precision-recall analyses confirm that SMOTE-based balancing significantly improves the robustness and reliability of multiclass intrusion detection in smart water IoT environments.

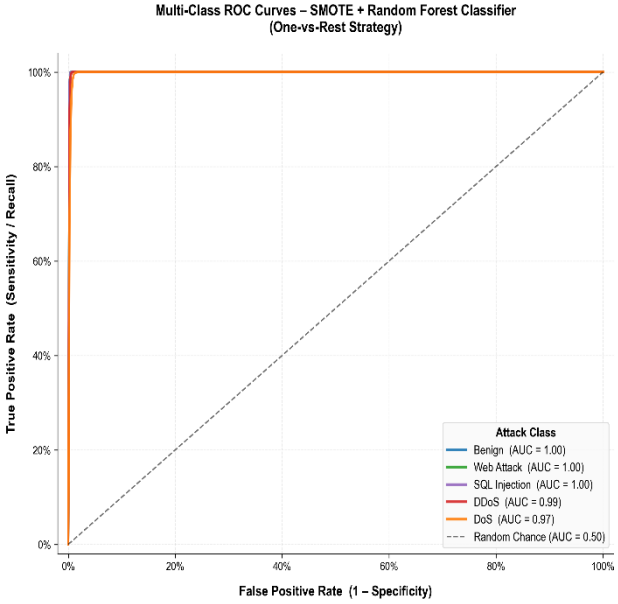

Figure 13. ROC curve for enhanced model

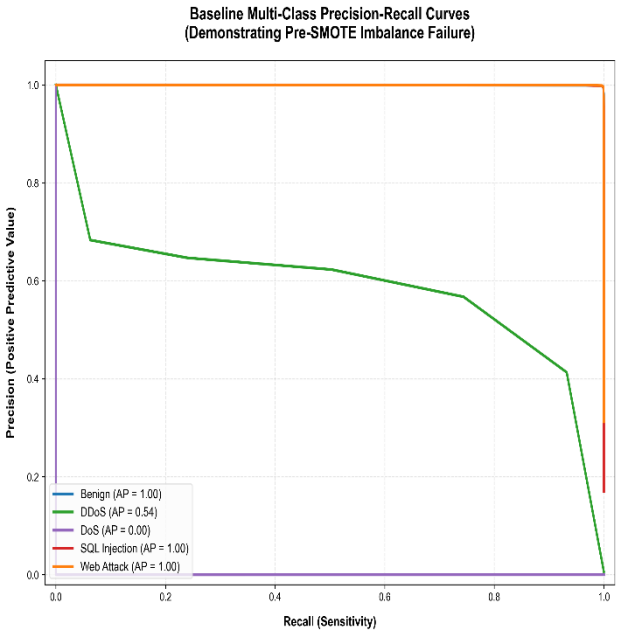

Figure 14. Precision-recall curve for baseline model

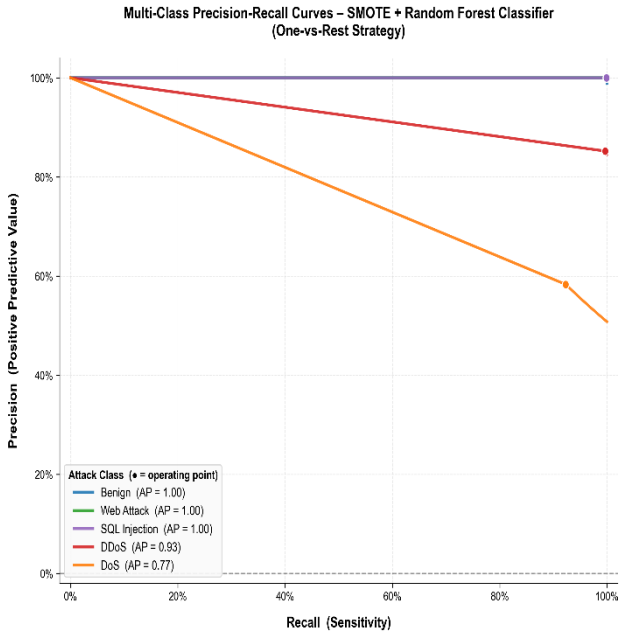

Figure 15. Precision-recall curve for enhanced model

6.4 Study limitations and future improvements

Despite the strong improvements achieved through SMOT-based balancing, several limitations remain. Synthetic oversampling may introduce artificial correlations among features and lead to overfitting, particularly for extremely small minority classes such as DoS. In addition, the Random Forest hyperparameters were intentionally fixed (n_estimators=100, random_state=42) to isolate the impact of class balancing, and extensive hyperparameter optimization was not performed. The current evaluation relies on a stratified split rather than temporal validation, limiting assessment under real-time streaming conditions. Future work should therefore investigate advanced resampling methods such as ADASYN and SMOTE-Tomek, apply automated hyperparameter tuning, and incorporate temporal as well as adversarial robustness evaluation to improve deployment reliability in smart water IoT environments.

7. Conclusion

The study tackles multiclass threat classification in real-time smart water IoT ecosystems, where class imbalance is widespread. Comparing an initial training regime based on the Random Forest with an SMOTE-rebalanced training regime, we can see that high aggregate accuracy can conceal operational failures in the presence of rare but significant attacks that were undetected. Our findings highlight that, without appropriate mitigation of imbalance, even complex ensemble classifiers do not necessarily detect important minority threats. The addition of SMOTE to the training pipeline addresses this drawback, producing an intrusion detection architecture that is practical to implement. The suggested method achieves an accuracy of 99.88% and a recall of 92.31% on DoS and 99.66% on DDoS, respectively, with a macro-F1 of 0.60 to 0.93. The study contributes not only a high-performing intrusion detection framework but also a reproducible experimental analysis demonstrating why class imbalance is essential for operational reliability in smart water cyber-physical systems. The suggested framework can be reinforced and developed in some directions in the future:

(1) Federated learning for privacy-preserving IoT analytics: Develop federated learning frameworks for smart water networks. Enable distributed model training across edge devices without sharing raw data.

(2) Edge Intelligence and TinyML: Design lightweight ML models (TinyML) for deployment on resource-constrained sensors. Enable real-time decision-making at the edge.

(3) Integration with smart city ecosystem: Explore the integration of smart water meter infrastructure within broader smart city ecosystems. Future research should focus on developing standardized communication protocols, secure data exchange mechanisms, and AI- driven cross-domain analytics to support seamless integration within smart city developments.

(4) Graph Neural Networks (GNNs) for Water Networks: Can apply GNNs for leak detection, flow prediction, and attack propagation analysis.

Reproducibility Statement

All experiments were performed on a 14-core CPU with 16GB RAM, in a hybrid Apache Spark-Pandas-Scikit-learn framework to guarantee computational reproducibility. Apache Spark was used to distribute preprocessing and feature engineering, and Pandas, Scikit-learn, and Imbalanced-learn libraries were used for anomaly scoring, resampling, and classification.

(1)Hyperparameters used in RF and SMOTE application

RandomForestClassifier(n_estimators=10,
warm_start=True,
random_state=42,
n_jobs=-1)
SMOTE(sampling_strategy='auto',k_neighbors=5,
random_state=42)

(2) Fixed random seeds(random_state=42) were constantly applied during train-test split, SMOTE balancing, Isolation Forest initialization, and Random Forest training to ensure deterministic and reproducible experiment behavior.

(3) Due to privacy and critical infrastructure security considerations, the dataset cannot be made publicly available. However, preprocessing scripts and anonymized samples are planned for release to support reproducibility and future research.

Ethical issue

The authors are aware of and comply with best practices in publication ethics, specifically regarding authorship (avoidance of guest authorship), dual submission, manipulation of figures, competing interests, and compliance with research ethics policies. The authors adhere to publication requirements that the submitted work is original and has not been published elsewhere.

Data availability statement

The manuscript contains all the data. However, additional data will be provided by the corresponding author upon reasonable request.

Conflict of interest

The authors declare no potential conflict of interest.

References

[1] M. M. Althobaiti. Intelligent intrusion detection for IoT and cyber-physical systems using machine learning. International Journal of Advances in Applied Sciences, vol. 14, no. 1, pp. 1–10,2025. DOI: 0.21833/ijaas.2025.06.009

[2] S. U. Jan; S. Ahmed; V. Shakho; I. Koo. Toward a lightweight Intrusion Detection System for the Internet of Things. IEEE Access, vol. 7,2019.DOI: 10.1109/ACCESS.2019.2907965

[3] T. Saranya; S. I. Priyadharshini. A Dual-strategy framework for cyber threat detection in imbalanced,high-dimensional data across heterogeneous networks. IEEE Access, vol. 13, pp. 125313-125331,2025. DOI: 0.1109/ACCESS.2025.3582788

[4] R. Ahmad; I. Alsmadi. Machine learning approaches to IoT security: A systematic literature review. Internet of Things,vol. 14,2024. DOI: 10.1016/j.iot.2021.100365

[5] L. Xiao. IoT security techniques based on machine learning: How do IoT devices use ai to enhance security? IEEE Signal Processing Magazine, vol.35,2019.DOI:10.14445/22312803/IJCTTV 7I2P110

[6] H. Fares; M. Zeroual; A. Karim; Y. Maleh; Y. Baddi; N. Aknin. Machine learning, deep learning and ensemble learning based approaches for intrusion detection enhancement. Edpacs, vol. 69, no. 1, pp.1–15,2024. DOI: 10.1080/07366981.2024.2422645

[7] A. Alrefaei; M. Ilyas. Using machine learning multiclass classification technique to detect IoT attacks in real time in Italian National ConferenceonSensors,pp.1-6,2024. DOI:10.3390/s24144516

[8] M. Bokhari; M. Z. Khan; F. Masoodi; M. Zeyauddin. Bagging ensemble model performance in IoT cyberattack detection: A comprehensive evaluation.In 12th International Conference on Computing for Sustainable Global Development (INDIACom),2025, DOI:10.23919/INDIACom66777.2025.1111 5395

[9] Y. Zhou; Guang Cheng; Shanqing Jiang; Mian Dai. Building an efficient intrusion detection system based on feature selection and ensemble classifier.Computer Networks,Vol.174,2020. DOI: 10.1016/j.comnet.2020.107247

[10] A. Thakkar; R. Lohiya. Attack classification of imbalanced intrusion data for IoT network using ensemble learning-based deep neural network. IEEE Internet of Things Journal, vol. 10, no.13, 2023.DOI:10.1109/JIOT.2023.3244810

[11] K. Shaukat; S. Luo; V. Varadharajan. Performance comparison and current challenges of using machine learning techniques in cyber-security. Energies, vol.13, no. 10, p. 2509,2020.DOI: 10.3390/en13102509

[12] R. Tanty; P. B. Dash; J. Nayak; B. Naik. Intelligent Intrusion Detection in military IoT networks using recursive feature elimination with extreme gradient boosting. In 2025 4th International Conference on Range Technology (ICORT), pp. 1–6,2025,.DOI: 10.1109/ICORT64008.2025.11115621

[13] D. Elreedy; A. F. Atiya. A comprehensive analysis of synthetic minority oversampling technique (SMOTE) for handling class imbalance. Information Sciences, vol. 505,2020.DOI: 10.1016/j.ins.2019.07.070

[14] C. Hazman. lids-sioel: Intrusion detection framework for IoT-based smart environments security using ensemble learning, Cluster Computing, vol. 26,2022.DOI: 10.1007/s10586-022-03810-0

[15] A. Abbas; M. A. Khan; S. Latif. A new ensemble-based intrusion detection system for internet of things. Arabian Journal for Science and Engineering,vol.47,no.2,2025. DOI: 10.1007/s13369-021-06086-5

[16] R. Chowdhury. An optimal feature-based network intrusion detection system using bagging ensemble method for real-time traffic analysis. Multimedia Tools and Applications,vol.81,no.28,2023.DOI: 10.1007/s11042-022-12330-

[17] S. S. Sonawane. Optimized deep feature analysis for enhanced botnet attack prediction in IoT networks. Journal of Information Systems Engineering&Management,vol.10,2025.DOI:10.52783 /jisem.v10i3s.413

[18] M. Douiba; S. Benkirane; A. Guezzaz; M. Azrour. An improved anomaly detection model for IoT security using decision tree and gradient boosting. The Journal of Supercomputing, vol. 79, no. 3, pp. 3392–3411, 2023.DOI:0.1007/s11227-022-04783-y

[19] L. Yang. Intrusion detection based on approximate information entropy for random forest classification. in 2019 4th International Conference on Big Data and Computing (ICBDC 2019), 2019.DOI: 10.1145/3335484.3335488

[20] N. Gupta; V. Jindal; P. Bedi. CSE-IDS: Using cost-sensitive deep learning and ensemble algorithms to handle class imbalance in network-based intrusion detection systems. Computer Security,vol.112,2022. DOI:10.1016/j.cose.2021.102499

[21] M. Altalhan; A. Algarni; M. T.-H. Alouane.Imbalanced data problem in machine learning: A Review. IEEEAccess,vol.13,2025. DOI: 10.1109/ACCESS.2025.3531662

[22] S. Bagui; K. Li. Resampling imbalanced data for network intrusion detection datasets. Journal of Big Data,vol. 8, no. 1,2021. DOI: 10.1186/s40537-020-00390-x

[23] M. A. Talukder; M. Khalid; N. Sultana. A hybrid machine learning model for intrusion detection in wireless sensor networks leveraging data balancing and dimensionality reduction. Scientific Reports, vol.15, no. 1,2025.DOI: 10.1038/s41598-025-87028-1

[24] S. Sadhwani; B. Manibalan; R. Muthalagu; P. Pawar. A lightweight model for DDoS attack detection using machine learning techniques, Applied Sciences, MDPI,

vol. 13, no. 15, p. 8712,2023.DOI: 10.3390/app13179937
[25] V. Patel; H. Shukla; A. Raval. Enhancing botnet detection with machine learning and explainable AI: A step towards trustworthy ai security. International Journal for Multidisciplinary Research,vol. 7, no. 1,2025.DOI: 10.36948/ijfmr.2025.v07i02.39353
[26] J. Manokaran; G. Vairavel; J. Vijaya. PPFCM-smote: a novel balancing system for anomaly detection in IoT edge using probabilistic possibilistic fuzzy clustering and smote. International Journal of Information Technology, 2024.DOI: 0.1007/s41870-024-02129-w
[27] M. B. Musthafa. Optimizing IoT intrusion detection using balanced class distribution, feature selection and ensemble machine learning techniques Sensors,MDPI, vol. 24, no. 13, 2024.DOI: 10.3390/s24134293
[28] Shaher Zyoud. Exploring the promising role of internet of things in urban water systems: a comprehensive global analysis of insights, trends and research priorities. Discover Internet of Things, vol.5,2025. DOI: 10.1007/s43926-025-00129-1
[29] W. Rajeh; M. M. Aborokbah; M. S., T. Alashoor; K. P. Tabnet-SFO: An Intrusion Detection Model for smart water management in smart cities. International Journal of Intelligent Systems, vol.2025,p.6281847,2025.DOI: 10.1155/int/6281847
[30] N. N. Tilakarathne and W. D. Madhuka Priyashan. An Overview of Security and Privacy in Smart Cities.EAI/Springer Innovations in Communication and Computing (2022): 21–44, DOI:10.1007/978-3-030-82715-1_2.
[31] O. D. Okey. Boostedenml: Efficient technique for detecting cyberattacks in IoT systems using boosted ensemble machine learning. Sensors, MDPI, vol. 22, no. 19,2022.DOI: 10.3390/s22197409
[32] P. K. Keserwani. A Smart Anomaly-based Intrusion Detection System for the Internet of Things (IoT)network using gwo-pso-rf model, Journal of Reliable Intelligent Environments, vol. 7, no. 1,2021.DOI:10.1007/s40860-020-00126-x
[33] A. S. Ahanger; S. M. Khan; F. Masoodi. Intrusion detection system for IoT environment using ensemble approaches. In 2023 10th International Conference on Computing for Sustainable Global Development (INDIACom), pp. 1–6,2023.URL:https://ieeexplore.ieee.org/abstract/document/10112382
[34] P. Verma. A novel intrusion detection approach using machine learning ensemble for IoT environments. Applied Sciences, vol. 11, no. 21,2021.DOI: 10.3390/app112110268
[35] Y. Cao. An intrusion detection system based on stacked ensemble learning for IoT network. Computers and Electrical Engineering,vol. 110,2023. DOI:10.1016/j.compeleceng.2023.108836
[36] W. Lian; Guoqing Nie; Bin Jia; Dandan Shi; Yongquan Liang. An intrusion detection method based on decision tree-recursive feature elimination in ensemble learning. Mathematical Problems in Engineering, vol 2020, issue 1,2020. DOI: 10.1155/2020/2835023
[37] M. B. Pranto; Md. Hasibul Alam Ratul; Md. Mahidur Rahman; Ishrat Jahan Diya; Zunayeed-Bin Zahir. Performance of machine learning techniques in anomaly detection with basic feature selection strategy-a network intrusion detection system. Journal of Advanced Information Technology, vol. 13, no. 1,2022. DOI: 10.12720/jait.13.1.36-44
[38] S. Alangari. An unsupervised machine learning algorithm for attack and anomaly detection in IoT sensors," Wireless Personal Communications,vol. 144,2024. DOI: 10.1007/s11277-023-10811-8
[39] M. N. Kanyama; F. B. Shava; A. M. Gamundani; A. Hartmann. Machine learning applications for anomaly detection in smart water metering networks: A systematic review. Physics and Chemistry of the Earth, vol. 134, p. 103558, 2025. DOI: 10.1016/j.pce.2024.103558
[40] J. Wang; J. E. van Zyl; L. Wen;Y. Li; S. Che. The impact of smart meter programmes on household water consumption: evidence from new Zealand, Journal of Behavioral and Experimental Economics, vol. 118, p. 102413, 2025. DOI: 10.1016/j.socec.2025.102413
[41] Hajar Hameed Addeen, Yang Xiao, Jiacheng li, and Mohsen Guizani. A Survey of Cyber-Physical Attacks and Detection Methods in Smart Water Distribution Systems. IEEE Access,vol-9,2021. DOI:10.1109/ACCESS.2021.3095713