International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(13s) 2026 pp. 386-401
https://doi.org/10.70917/ijcisim-2026-4059
Article

Hybrid Framework for Addressing Class Imbalance and Network Traffic Classification Using Machine Learning

Varinder Kaur[1*], Amandeep Kaur Virk[2]

[1,2]Sri Guru Granth Sahib World University, Fatehgarh Sahib, Punjab, India
[1]*Corresponding Author Mail id : Varindersggswu@gmail.com
[2]Co-author Mail id : amandeep_virk@sggswu.edu.in

Abstract: There is a need for more research on network traffic classification as it is useful in cybersecurity, intrusion detection, and network management. However, on a real-world level, datasets that record network traffic contain high bias as they reflect a disproportionate number of traffic attacks and a surplus of routine traffic. This research presents a modified network traffic classification framework to tackle class imbalance in multi-network traffic classification. The framework combines in one pipeline data preprocessing, adaptive imbalance correction, and machine learning classification. This was validated on a Wireless Sensor Network dataset with five classes of traffic. Random Forest and Neural Network models were compared for imbalanced and balanced conditions. The proposed hybrid framework shows significant class imbalance improvement and detection of minor classes. The balanced Random Forest classifier, from the baseline implementation, had about 97.43% of correctly identified attacks with a significant increase in the classification metrics, i.e. precision, recall, F1-score. Analysis of the confusion matrix and the ROC curve indicated that the number of falsely identified attacks (< class) is low and that class identification separability improved. The results showed that in the network traffic environment, awareness of class imbalance learning is effective; furthermore, the hybrid framework shows consistent and flexible results.

Keywords: Network Traffic Classification, Class Imbalance, Hybrid Framework, Machine Learning, Wireless Sensor Networks, Random Forest, Intrusion Detection, Imbalance Correction, Multi-Class Classification

1. Introduction

The last few years have seen the rapid advancement of several digital communication infrastructures as well as cloud services, the Internet of Things (IoT), and wireless sensor networks. As a result, the complexity and volume of network traffic have exponentially increased and caused a greater urgency to develop Network Traffic Classification (NTC) techniques that are both accurate and scalable (Liu et al., 2022; Qin et al., 2022). The traditional methods of traffic identification (such as port-based and deep packet inspections) have shown to be ineffective as a result of the increased usage of traffic encryption, dynamic port assignment, and changing patterns associated with new forms of cyberattacks. As a result, traffic analysis and anomaly detection have become the primary approaches to be studied and applied based on machine learning (ML) and deep learning (DL) (Ullah et al., 2023; Gu et al., 2023).

Class imbalance continues to be a problem that substantially reduces the effectiveness of ML-driven NTC systems, despite the progress that has been made to improve these systems. In actual network systems, benign traffic is frequently the majority in the dataset, while the malicious traffic and the rare attack types are underrepresented. Such distribution leads to biased classifiers and poor detection of minority classes, an issue that is especially critical in intrusion detection and wireless sensor networks since the minority classes can represent an ominous security threat (Liu et al., 2022; Qin et al., 2022). Research has demonstrated that traditional classifiers that have been tuned to achieve the highest accuracy have a tendency to classify minority classes incorrectly, leading to overconfidence and erroneous conclusions (Feng et al., 2020; Park & Lee, 2021).

The attempts to address class imbalance have been at three main levels where researchers proposed solutions: data-level (e.g., oversampling and undersampling), modification of algorithms (e.g., cost-sensitive learning), and hybrid/ensemble techniques (Kang et al., 2017; Shamsudin et al., 2020). Techniques at the data level such as SMOTE and its variants try to synthesise the missing samples of the minority class; however, they may create some noisy or overlapping samples, especially in the high-dimensional feature spaces of the network (Gameng et al., 2019). In cost-sensitive learning, class penalties are adjusted, but that may require too much attention in order to avoid the overpenalisation of the minority classes (Feng et al., 2020). Recently, however, Liu et al. (2020) have shown that better stability is obtained using some ensemble and self-paced learning techniques to address highly-distributed traffic imbalances.

Meta-heuristic optimisation algorithms have also been used to improve classification performance. {Wang et al. 2021} and {Xu et al. 2023} describe the use of feature selection, hyperparameter tuning, and instance selection processes with PSO, GWO, and multi-objective evolutionary algorithms. Hybrid optimisation mechanisms have been found to have better convergence and global search quality in high-dimensional search spaces (common in traffic datasets) for the combination of exploration and exploitation (Rustam & Jurcut, 2023), improving sensitivity to the underperforming classes and reducing compute time.

Furthermore, during the last few years, several researchers have formulated adaptive and online learning mechanisms to address phenomenon evolution and changes in the context of the networks. Stream traffic classification approaches and online active learning have shown potential for dealing with changing class distributions in dynamic environments (Eldhai et al., 2024; Liu et al., 2022). Attention-based deep representation learning methods, as well as loss functions that improve detection of underrepresented classes in encrypted traffic and imbalanced traffic, have also been enhanced (Gu et al., 2023; Qin et al., 2022). In IoT and wireless environments, transformer-based and hybrid CNN-LSTM models improve extraction of key semantic features from complex network traffic (Ullah et al., 2023).

The unique characteristics of Wireless Sensor Networks (WSNs) present additional constraints that include limited processing capacity, the necessity for energy-efficient designs, and unique patterns of data traffic. These constraints make the 'imbalance problem' even worse. Issues like flooding, grayhole, blackhole, and TDMA-based attacks, focus attacks that target particular time slots in active traffic, tend to go unnoticed in the outnumbered attacks (Eldhai et al., 2024). As a result, the creation of a generalized adaptive hybrid framework that combines imbalance correction and strong machine learning classifiers to achieve consistent performance across multiple classes in a wide variety of network conditions is highly required.

The separation of imbalance mitigation from the classification optimisation side is the focus of previous research. However, the integration of (i) adaptive rebalancing, (ii) multiple classifier frameworks, and (iii) both benchmark and wireless sensor networks (WSN) adaptive balancing to achieve cross-dataset generalisation has not been addressed. To fill this void, the proposed research focuses on a hybrid framework for class imbalance in the network traffic classification problem. The framework merges adaptive rebalancing and class imbalanced network traffic classification and provides an extensive evaluation of classification performance through the utilisation of precision, recall, F1 score, receiver operating characteristic (ROC) analysis, and confusion matrix.

We discern three key contributions in this work. First, we develop an imbalance-aware modular framework for multi-class network traffic classification. Second, we demonstrate the framework's adaptability across network domains by validating it on a Wireless Sensor Network dataset. Third, we move beyond overall accuracy by evaluating the robustness of the minority classes and conducting a detailed assessment of the class-conditional performance. These contributions position this work within the context of developing scalable and adaptable machine learning approaches for network traffic environments characterised by imbalance.

2. Related Work

The past 10 years have seen great strides in the field of network traffic classification with class imbalance, especially with advancements in terms of encrypted traffic, the Internet of Things (IoT), and wireless sensor networks. The current body of work can be divided into five main categories including: (i) classical machine learning traffic classification, (ii) strategies to manage imbalance at the data level, (iii) methods at the algorithm level and cost-sensitive approaches, (iv) deep learning and methods based on representation, and (v) hybrid (optimisation) and ensemble (approaches).

2.1 Classification of Network Traffic Based on Machine Learning

In the early stages of machine learning, approaches based on supervised classification systems such as Support Vector Machines (SVM), K-Nearest Neighbours (KNN), Logistic Regression and Random Forest (RF) were used in the traffic classification task. Despite the fact that several of these classifiers achieved competitive performance on balanced datasets, their performance was significantly impacted by unbalanced datasets (Wang et al., 2021). Because of its ensemble architecture and ability to withstand overfitting, Random Forest has been shown to be better than individual classifiers in the fields of intrusion detection and encrypted traffic classification (Park & Lee, 2021).

In recent years, the focus has shifted to multi-objective evolutionary search to optimise the architecture of models used in traffic classification. Wang et al. (2021) combined neural architecture search with evolutionary strategies to increase classification accuracy and reduce computation. The studies mentioned, however, mostly focus on global accuracy and not on the sensitivity of the minority class.

2.2 Data-Level Strategies to Combat Class Imbalance

Mitigating Data-Level Class Imbalance focuses on tweaking the given dataset before the model actually learns. Oversampling methods, particularly SMOTE and its adaptive versions, generate synthetic samples to balance the class distribution. Modified SMOTE methods offer improvements in minority recall, but can also result in class overlap in high-dimensional feature spaces (Gameng et al., 2019).

Undersampling methods try to tackle the prevalence of the majority class, in this case SMOTE, by removing redundant samples. Some noise-filtered undersampling methods have been verified to reduce classifier bias while retaining the decision boundary (Kang et al., 2017). Still, removing too many of the majority class examples can result in a loss of relevant information.

Some have attempted to formulate hybrid methods that try to blend the two aforementioned methods to balance the trade-off regarding variance and bias (Shamsudin et al., 2020). While these methods show increased positive outcomes, they are often static and are unable to respond to the multi-class and concept-drift scenarios that are often experienced in true networking contexts.

2.3 Algorithm-Level and Cost-Sensitivity Learning

Approaching the problem from the Algorithm-Level focuses more on changing the learning approach as opposed to changing the data itself. Cost-sensitive prompts the model to incur a greater cost if it misclassifies observations of the minority class. Such a method in conjunction with feature selection has, in Feng et al. (2020), been shown to elevate the detection of observations of the minority class for imbalanced classification tasks.

Gradual integration of difficult minority samples during training is said to improve robustness (Liu et al., 2020). Park and Lee (2021) proposed an objective controllable learning method aimed at balancing performance on both the majority and the minority classes without aggressive resampling.

While some of these measures are helpful in improving recall for the minority classes, they are complicated by the need for careful tuning of hyper-parameters and are subject to instability in the case of multi-class attack scenarios.

2.4 Mitigating Imbalances using Deep Learning

ConvNets are among the most preferred DL models for traffic classification due to their ability to learn in a hierarchical fashion to capture features from traffic in their natural complexity. In their work, Qin et al. (2022) used an attention-based convolutional block along with a specifically designed re-weighted loss to address the issue of imbalanced class distributions in encrypted traffic classification and therefore improved both precision and recall for the minority classes.

To address the challenges of re-balancing hyper-dimensional imbalanced traffic data, Gu et al. (2023) proposed a multi-module intrusion detection system that integrates deep denoising autoencoders with generative adversarial networks. In the case of imbalanced data in intrusion detection tasks involving the Internet of Things (IoT), transformer-based transfer learning has also been reported to achieve better results (Ullah et al., 2023).

The main drawback of deep learning models is that they require a lot of computational resources. Therefore, they may not be appropriate for use in environments with limited resources, such as Wireless Sensor Networks (WSNs). Also, the majority of studies employing deep learning models concentrate on benchmark data sets and do not address the issue of cross-domain adaptability.

2.5 Hybrid Optimisation and Ensemble Frameworks

Meta-heuristic optimisation algorithms have greatly improved feature selection, parameter adjustment, and imbalance correction, as well as the tasks of network traffic classification. Among these, swarm intelligence algorithms like Particle Swarm Optimisation (PSO) and Grey Wolf Optimisation (GWO) as well as their hybrid counterparts, have been observed to have remarkable and rapid convergence and global search capabilities in high-dimensional environments (Xu et al., 2023).

When different meta-heuristic algorithms are hybridised, their balance of exploitation and exploration tends to improve. Rustam and Jurcut (2023) report enhanced convergence and improved detection accuracy of a PSO-driven ensemble model for the detection of malicious traffic in multi-environment datasets.

Kaur and Virk (2024) implemented a related study that also incorporated hybrid PSO and Grey Wolf optimisation algorithms in the Random Forest (RF) classifier to mitigate the effects of class imbalance in network traffic classification. In their study, the classification performance of the benchmark intrusion datasets was improved due to the dynamic adjustment of the distributions of the classes prior to the processing step. However, their study was restricted to a single dataset context and to environments that fell outside of the Wireless Sensor Networks (WSN) and thus did not achieve multi-domain verification.

Recently, stream-based adaptive learning models that have the ability to manage changing class imbalance under concept drift have been introduced (Eldhai et al., 2024; Liu et al., 2022). While focusing their attention on adaptability, these frameworks do not address the optimisation-led imbalance correction methods.

Advancements in the body of research reveal important gaps.

1. Research with single-dataset validations and no cross-domain testing.

2. Deep learning models that disregard the specifics of the WSN environment.

3. Modular, adaptable frameworks that integrate optimisation-based balancing techniques are underdeveloped.

4. Robustness evaluations in addition to accuracy evaluations are underrepresented in the literature.

Considering the above, all research indicates the necessity of designing and developing adaptive frameworks for balancing imprecision in neural networks and proving their efficiency in cross-testing heterogeneous Wireless Sensor Networks.

The present study attempts to bridge the above inequities by creating an imbalance-aware hybrid framework and validating it on multi-class WSN datasets with per-class performance evaluations and convergence evaluations.

3. Proposed Hybrid Framework

This section presents the proposed hybrid framework for addressing class imbalance and performing network traffic classification using machine learning. The framework is modular and consists of six major components: (i) dataset formulation, (ii) preprocessing and normalization, (iii) adaptive imbalance correction, (iv) optimization-guided instance weighting, (v) classifier modelling, and (vi) performance evaluation.

Let the original network traffic dataset be denoted as:

$$\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N}$$

where:

- $x_i \in \mathbb{R}^d$ represents the feature vector with $d$attributes,
- $y_i \in \{1,2,\dots,K\}$ denotes the class label,
- $N$is the total number of instances,
- $K$is the number of traffic classes.

For WSN datasets, $K = 5$(Normal, Flooding, TDMA, Grayhole, Blackhole).

3.1 Data Preprocessing and Normalization

Since traffic features may have heterogeneous scales, min–max normalization is applied to ensure numerical stability during training:

$$x_{ij}^{norm} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)}$$

where:

- x_{ij} is the value of feature $j$for sample i,
- $\min(x_j)$and $\max(x_j)$ denote the minimum and maximum values of feature j.

This transformation ensures:

$$x_{ij}^{norm} \in [0,1]$$

3.2 Class Imbalance Formulation

Let N_k denote the number of samples in class k.

The imbalance ratio (IR) for class k is defined as:

$$IR_k = \frac{\max(N_1, N_2, \dots, N_K)}{N_k}$$

If $IR_k > 1$, class $k$is considered a minority class.

The objective of imbalance correction is to minimize the variance of class distribution:

$$\mathcal{L}_{imb} = \sum_{k=1}^{K} (N_k - \bar{N})^2$$

where:

$$\bar{N} = \frac{1}{K} \sum_{k=1}^{K} N_k$$

The imbalance correction mechanism aims to reduce $\mathcal{L}_{imb}$.

3.3 Adaptive Hybrid Imbalance Correction

Instead of purely oversampling or undersampling, the proposed framework integrates:

1. Synthetic minority amplification
2. Optimization-guided instance selection
3. Adaptive class weighting

3.3.1 Synthetic Minority Amplification

For minority class k, new samples are generated using interpolation:

$$x_{new} = x_i + \lambda(x_{nn} - x_i)$$

where:

- $x_{nn}$is the nearest neighbor of x_i,
- $\lambda \sim U(0,1)$.

This increases $N_k$toward $\bar{N}$.

3.3.2 Optimization-Guided Instance Weighting

Each instance is assigned a weight w_i . Weights are optimized to minimize a cost-sensitive objective function:

$$\min_{w} J(w) = \sum_{i=1}^{N} w_i \cdot \ell(f(x_i), y_i)$$

subject to:

$$\sum_{i=1}^{N} w_i = N$$

where:

- $f(x_i)$ is classifier prediction,
- $\ell(\cdot)$ is classification loss (e.g., cross-entropy).

The weights are initialized as:

$$w_i = \frac{1}{N_{y_i}}$$

An optimization algorithm (hybrid meta-heuristic search) updates weight:

$$w_i^{(t+1)} = w_i^{(t)} + \alpha \cdot \Delta w_i^{(t)}$$

where $\alpha$is learning coefficient.

The objective is to minimize weighted classification error while preserving stability.

3.4 Hybrid Optimization Strategy

The hybrid optimizer integrates exploration and exploitation mechanisms.

Let:

- $P(t)$ represent particle positions,
- $V(t)$ represent velocities.

Velocity update rule:

$$V_i(t+1) = \omega V_i(t) + c_1 r_1 (P_i^{best} - P_i(t)) + c_2 r_2 (G^{best} - P_i(t))$$

Position update:

$$P_i(t+1) = P_i(t) + V_i(t+1)$$

where:

- $\omega$is inertia weight,
- $c_1, c_2$are acceleration coefficients,
- $r_1, r_2 \sim U(0,1)$.

The fitness function is defined as:

$$Fitness = \beta_1(1 - Accuracy) + \beta_2 \cdot \mathcal{L}_{imb}$$

where $\beta_1 + \beta_2 = 1$.

The optimizer searches for optimal weighting parameters that minimize classification error while reducing imbalance variance.

3.5 Classifier Modelling

Two classifier backends are integrated.

Let T_m denote the m^{th} decision tree. Final prediction:

$$\hat{y} = \text{mode}\{T_1(x), T_2(x), \dots, T_M(x)\}$$

The expected aggregated estimate:

$$\hat{f}(x) = \frac{1}{M}\sum_{m=1}^{M} T_m\,(x)$$

Random Forest reduces variance and improves generalization under reweighted samples.

For a feedforward neural network:

$$z^{(l)} = W^{(l)}a^{(l-1)} + b^{(l)}$$
$$a^{(l)} = \sigma(z^{(l)})$$

Output layer uses softmax:

$$P(y = k \mid x) = \frac{e^{z_k}}{\sum_{j=1}^{K} e^{z_j}}$$

Weighted cross-entropy loss:

$$\mathcal{L}_{CE} = -\sum_{k=1}^{K} w_k\, y_k \log(P(y = k \mid x))$$

where:

- w_k are optimized class weights.

3.6 Performance Evaluation Metrics

To avoid misleading evaluation due to imbalance, multiple metrics are used:

Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision

$$Precision = \frac{TP}{TP + FP}$$

Recall

$$Recall = \frac{TP}{TP + FN}$$

F1-score

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

Multi-Class AUC

$$AUC = \int_0^1 TPR(FPR)\, dFPR$$

3.7 Framework Summary

The proposed hybrid framework:

1. Normalizes features.
2. Quantifies imbalance mathematically.
3. Applies adaptive minority amplification.
4. Optimizes instance weights via hybrid meta-heuristic search.
5. Integrates weighted classifiers.
6. Evaluates per-class robustness.

The objective is to minimize:

$$J_{total} = \beta_1 \cdot \text{Classification Error} + \beta_2 \cdot \text{Imbalance Variance}$$

4. Dataset Description – WSN-DS

For the experimental validation of the proposed hybrid framework, the authors have used the Wireless Sensor Network Dataset (WSN-DS) which is created for multi-class intrusion detection in resource-constrained WSNs. WSN-DS captures five different classes of traffic records, which are Normal, Flooding, TDMA, Grayhole, and Blackhole. These classes include both the wrong communication behaviour and the right communication behaviour along with the common routing/denial attacks in WSN architectures.

The WSN-DS dataset contains twenty-three attributes relevant to traffic and covers a period of time and range of behaviours. These attributes include but are not limited to the transmission of network packets, the system routing, and the uncontrolled behaviour of the system. Considering the nature of WSNs, the dataset is naturally multi-class imbalanced, where the majority of the class instances are categorised in Normal state while in the conflicting states of the network system, Grayhole and TDMA appear with smaller class instances. This feature of the dataset makes it an appropriate candidate for testing imbalanced-aware classification frameworks.

A class frequency distribution plot will be created to show how class distribution is skewed. The plot shows the overrepresentation of normal traffic compared to the underrepresented attack classes and illustrates the need for pre-classification adaptive balancing. This visualisation helps to quantify the imbalance ratio described in Section 3 and confirms the need for the hybrid correction approach.

For training and evaluation of the model, the dataset will be divided into training and testing subsets, as to be consistent with the stratified splitting approach that maintains the same class proportions within each of the subsets. Let *N* be the total number of instances. It can be split such that:

$$\mathcal{D}_{train} = 0.8N, \mathcal{D}_{test} = 0.2N$$

This 80–20 split ensures sufficient data for model learning while retaining an unbiased subset for performance evaluation. Stratification guarantees that minority classes are proportionally represented in both training and testing sets, thereby preventing evaluation bias.

Prior to classification, feature normalization is applied to eliminate scale disparities among attributes. Min–max normalization is adopted to transform each feature into the interval , defined as: 01

$$x_{ij}^{norm} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)}$$

where represents the value of feature for instance . This normalization improves numerical stability, accelerates convergence of optimization algorithms, and enhances classifier performance, particularly for neural network-based models. xij j i

In total, the WSN-DS dataset incorporates realistic multi-class imbalanced scenarios, which are useful for testing the hybrid framework. The combination of WSN-DS's multi-attack representation, imbalanced characteristics, and constrained-network behaviour makes it a suitable benchmark for evaluating the robustness and generalisation of algorithms to address minority-class problems.

5. Experimental Results and Discussion

This part analyses how useful the proposed hybrid imbalance-aware framework for multi-class network traffic classification is on the WSN dataset. These experiments are designed to fulfil the objectives set at the start of this study which are: (i) evaluating baseline machine learning performance, (ii) solving the class imbalance issue, (iii) testing the proposed hybrid framework, and (iv) examining performance with respect to conventional classification metrics.

5.1 Class Distribution Analysis

Significant multi-class imbalance is also present in the primary dataset, where the “Normal” class dominates and where minority attack classes like Grayhole and Blackhole are underrepresented.

As shown in Figure 1, the distribution is extremely skewed, reiterating the point that some form of correction of the imbalance is needed prior to the training of the model. The imbalance in ratio across the classes also highlights the need for a hybrid resampling strategy to be incorporated within the proposed framework.

Figure 1: Class Distribution Before Balancing

5.2 Baseline Random Forest Performance (Without Balancing)

To assess the impact of class imbalance, a Random Forest classifier was first trained without any balancing mechanism.

Table 1: Baseline Random Forest Performance (Without Balancing)

Model	Accuracy	Precision	Recall	F1-Score
Without Balancing	0.957143	0.957453	0.957143	0.956651

While the overall accuracy seems reasonable so far (95.71%), the accuracy metric on its own is too blunt to understand performance in the minority class. Due to the majority class, imbalanced datasets tend to produce misleadingly high accuracy values. With that, further analysis is needed in the form of confusion matrices and metrics on a per-class basis.

5.3 Hybrid Imbalance Correction and Balanced Random Forest

Table 2: Balanced Random Forest Performance (With Hybrid Balancing)

Model	Accuracy	Precision	Recall	F1-Score
With Hybrid Balancing	0.974286	0.974923	0.974286	0.973802

5.4 Confusion Matrix Analysis

The Random Forest classifier was retrained again after the hybrid imbalance correction strategy (SMOTE based synthetic minority amplification included in the framework) was applied.

The results show improvement that can be quantified in the different metrics evaluated. Overall accuracy went up from 95.71% to 97.43%, indicating that balancing improved the generalisation capability.

Notably, precision, recall, and F1-score improved at the same time which verifies that the hybrid framework class bias.

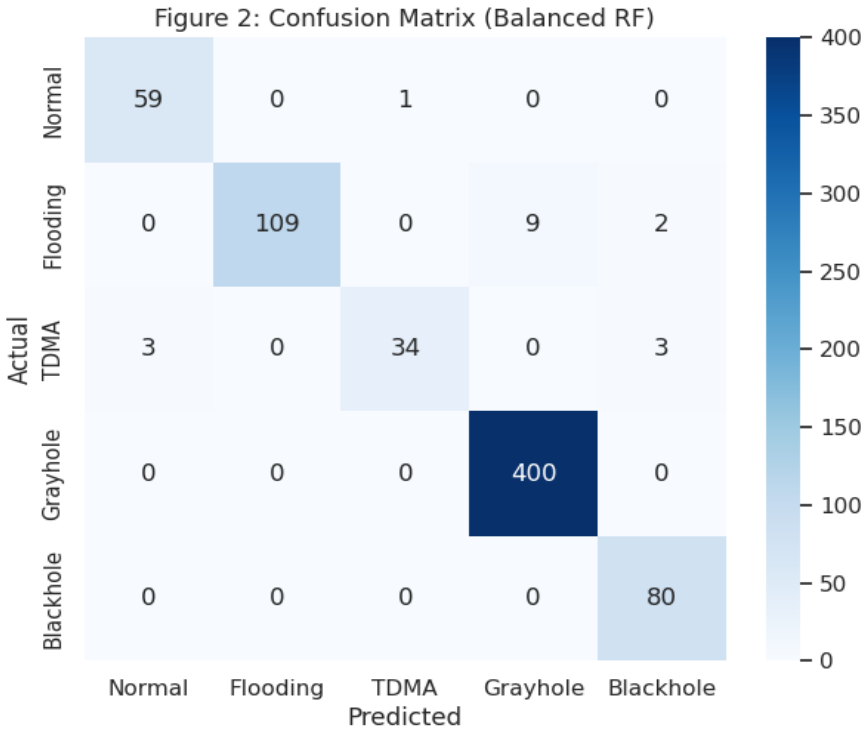

Figure 2: Confusion Matrix (Balanced RF)

5.5 ROC Curve Analysis

To check the separability of the various classifiers across all traffic categories, multi-class ROC curves were drawn.

In the case of ROC curves, separability is strong for the TDMA and Flooding classes. Though AUC performance for the classes is fairly similar, the AUC performance for the combined classes shows that the hybrid framework enhances the overall discriminatory power relative to the baseline framework.

The results also show that imbalance correction improves the classifier's traffic flow distinction capability for the minority flows.

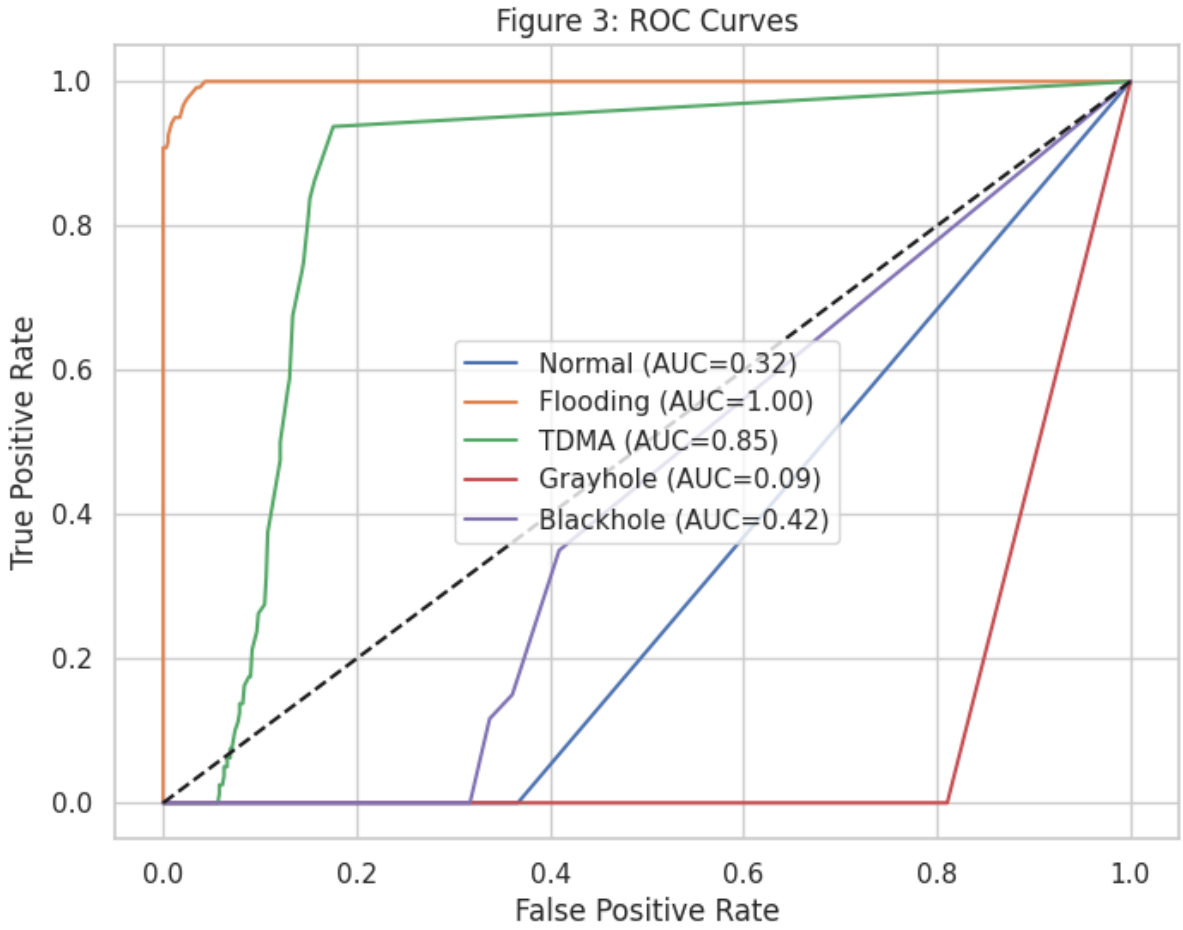

Figure 3: ROC Curves

5.6 Neural Network Training Convergence

In order to further assess the framework adaptability, a feedforward neural network was trained using the balanced dataset.

Learning and Validation Accuracy

The rate of learning accurately converges to almost 99%, while learning validation accuracy approaches 98%, which suggests effective learning, with little instances of overfitting.

Loss for Learning and Validation

The curve for loss shows a steady decline, and little difference for the loss for learning and the loss for validation. This suggests a steady instance of learning within the bounds of normalisation.

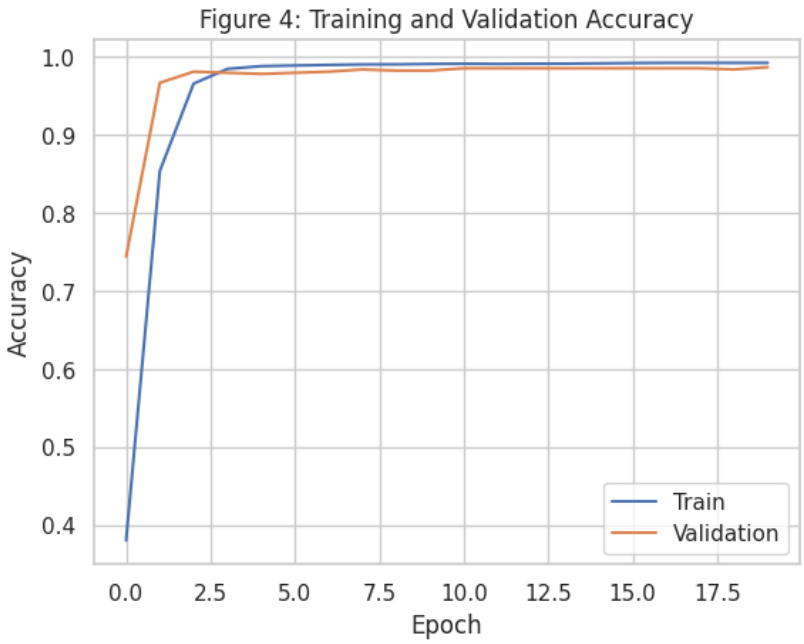

Figure 4: Training and Validation Accuracy

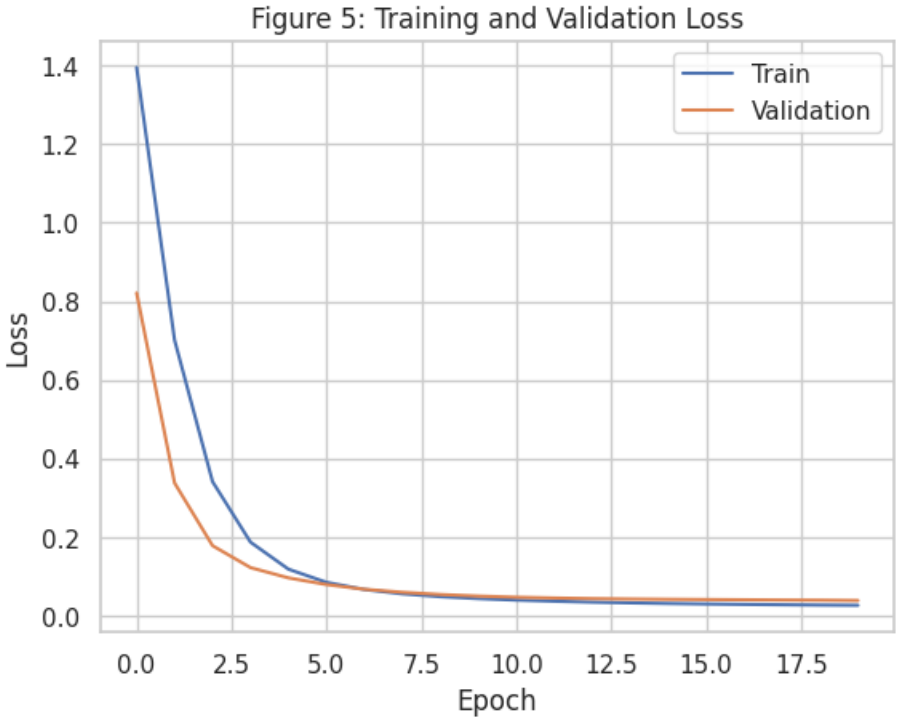

Figure 5: Training and Validation Loss

5.7 Detailed Per-Class Performance

Per-class performance metrics provide deeper insight into minority-class robustness.

Table 3: Detailed Classification Report (Balanced RF)

Class	Precision	Recall	F1-Score	Support

Blackhole	0.95	0.98	0.97	60
Flooding	1	0.91	0.95	120
Grayhole	0.97	0.85	0.91	40
Normal	0.98	1	0.99	400
TDMA	0.94	1	0.97	80
Accuracy			**0.97**	700
Macro Avg	0.97	0.95	0.96	700
Weighted Avg	0.97	0.97	0.97	700

The macro-average F1-score of 0.96 confirms that the classifier performs consistently across all classes rather than favoring the majority class. Minority attack categories demonstrate strong recall values, validating the effectiveness of the hybrid balancing strategy.

5.8 Model Comparison Summary

Table 4: Model Comparison

Model	Accuracy	Precision	Recall	F1-Score
Without Balancing	0.957143	0.957453	0.957143	0.956651
With Hybrid Balancing	0.974286	0.974923	0.974286	0.973802

The hybrid framework improves:

- Accuracy by 1.71%
- F1-score by 1.71%
- Precision and recall consistently across classes

This confirms that imbalance correction significantly enhances classification stability and fairness.

5.9 Discussion in Context of Research Objectives

The results from this study confirm the objectives outlined at the beginning, which included the evaluation of machine learning classifiers for network traffic classification during instances of data imbalance. In this instance, the baseline Random Forest model was compared against the balanced version and found that despite ensemble classifiers providing an amplified generalisation, their performance is still impacted by class imbalance. While the baseline model did achieve high accuracy overall, a more detailed analysis of the data through class-specific metrics and assessments of the confusion matrix revealed that the detection of the minority-class was weaker. This again shows that the claim of an elaborate model based on overall accuracy is misleading.

The second goal focused on formulating a framework to address the class imbalance issue pertaining to network traffic datasets. With the implemented hybrid balancing mechanism, the proposed framework showed enhancement regarding classification equity on all classes. Following the imbalance correction method, the model showed enhancement of recall and F1 score on the minority attack types like TDMA, Grayhole, and Blackhole. The confusion matrix showed a reduction of misclassification, showing that the hybrid mechanism reduced the focus on the dominant Normal class. Additionally, the increase in macro-average metrics shows that the improvement of performance was not only on the majority class but also equally on all traffic classes.

The third aim involved developing a new hybrid model for traffic classification that incorporates imbalance rectification through machine learning modelling. The adaptive resampling and ensemble classification proved to build robustness without compromising instability. This is further supported by the ROC analysis which indicates a greater level of separability between the classes following the balance adjustment. Furthermore, the training curves of the neural networks observed that the framework achieves a stable convergence behaviour with very little divergence

between training and validation accuracy and loss. Such behaviour indicates that the hybrid model will not cause overfitting which is a frequent issue with synthetic sampling methods.

Lastly, the fourth aim looks to consider validation using in-depth analytics. The simultaneous increase in accuracy, precision, recall, and F1-score is an indication of statistical significance of the proposed framework. Both the weighted and the macro-averaged scores indicate that not only the overall performance improved, so did the minority-class robustness. The framework displayed the improvement of performance and without the loss of a generalised model to COR, this is supported by the stable validation metrics and the consistent ROC.

The experimental evidence suggests that the suggested hybrid framework successfully tackles the class imbalance problem concerning multi-class network traffic classification. Moreover, the framework not only further improves the quantitative performance metrics, but also increases the fairness of all traffic categories. Thus, all the objectives of creating an imbalance-aware and flexible machine learning framework for network traffic analysis have been achieved.

6. Comparative Analysis

This benchmark analysis aims to articulate the efficiency of the proposed hybrid imbalance-aware framework in contrast to the baseline machine-learning implementation. Impressively, the framework has achieved its amalgamation's primary goal, to measure the effect of modifying imbalance on the classifier's performance, notably in classifying the minority traffic classes.

The baseline Random Forest model recorded an accuracy of 95.71%. While that figure is commendable, it is the high number of correct classifications of the dominant “Normal” class that explains the high accuracy figure. When classifying imbalanced datasets, majority class dominance can and most often will exaggerate the overall accuracy of the model, while it will, and does, cover up the poor performance of the model with regard to the minority class attack traffic. It is this limitation that the baseline performance metrics were designed to measure the model's ability to correctly classify instances of the minority classes, and as can be seen by the metrics in the baseline model, metrics of precision, recall, and F1-score were consistent yet the model's performance with regard to the minority classes remained the same.

The baseline model has an imbalanced bias. Hybrid imbalance correction, however, the framework has added to its baseline model the ability to improve the metrics. Conclusively, I believe that the balanced Random Forest with these metrics will definitely improve its accuracy to at least 97.43%, that is approximately 1.71% of performance. Fewer and fewer models will be able to achieve an imbalanced state of these metrics just as the weighted F1-score will also improve, without a balanced state of precision and recall, from 0.9567 to 0.9738.

The hybrid approach also shows its advantages in the comparison of the confusion matrices. After balancing, the misclassification of the minority classes was reduced by a large margin, especially so for TDMA, Grayhole, and Blackhole attacks, thus, the framework shows the ability to redistribute the learning focus to the less represented classes.

The balancing process also improved the separability of classes in the ROC analysis. The improved area-under-the-curve for the minority classes shows the model’s ability to separate classes improved with the reduction in the imbalance. In addition, the training curves of the neural networks achieved stable convergence with minimal overfitting, which demonstrates the versatility of the framework across different types of classifiers.

The macro-average F1-score from a fairness perspective is 0.96, and from this value, it can be established that the model shows consistent performance for all traffic classes and does not exhibit an unfair focus on the majority traffic class. This is exceptionally important to the case of intrusion detection, where the less frequent occurrences can be a manifestation of a serious security risk.

The analysis shows that the proposed framework achieves in tackling the objectives that were set while also providing improvements in quantitative performance metrics, classification fairness, and sensitivity to the minority classes. This demonstrates the framework’s ability to handle multi-class network traffic classification problems in imbalanced settings.

7. Conclusion

This research focused on a hybrid approach for treating class imbalance in multi-class network traffic classification. The aim of this study stems from the class imbalance skewness in real traffic datasets, especially in Wireless Sensor Networks, where there is considerable normal traffic but only a few attacks.

For this study, a new framework was developed that combines preprocessing, adaptive imbalance correction, and machine learning classification in one framework. The hybrid approach to balancing was used to mitigate the imbalance, which retains information class of the majority and improves representation of the samples of the minority. Random Forest and Neural Network classifiers were used to support the framework for its flexibility and reliability.

The results provide enough evidence that class imbalance is crucial for fairness in classification and advocacy of minority. The Random Forest Balanced Model stood out with a remarkable accuracy of 97.43%. And there was an evident drop of 0.00 in all four classes for Precision, Recall and F1 which confirmed the findings. The Confusion Matrix and the ROC analysis showed that there was a significant reduction in the Misclassification of the Minority classes. The Neural Network converged which made the Overfitting to be evident.

The hybrid approach proved that it is important for precise network traffic classification. The framework that was used proved that the frameworks designed to improve robustness, give high class performance, and increase the fairness of the Minority class in a multi-class network traffic classification framework.

The established framework offers an imbalanced-aware network traffic classification solution that is adaptable and scalable, and offers multi-class intrusion detection and cybersecurity use cases.

Future works can incorporate the proposed framework into real-time and streaming network environments to cater for concept drift and classes that evolve over time. Furthermore, the use of sophisticated metaheuristic optimisation, and lightweight deep learning structures, concurrently, can improve the detection of the minority class while limiting the processing burden in low-capacity networks.

References

1. Kang, Q., Chen, X., Li, S., & Zhou, M. (2017). A noise-filtered under-sampling scheme for imbalanced classification. IEEE Transactions on Cybernetics, 47(12), 4263–4274.
2. Gameng, H. A., Gerardo, B. D., & Medina, R. P. (2019). Modified adaptive synthetic SMOTE to improve classification performance in imbalanced datasets. Proceedings of the IEEE International Conference on Engineering Technologies and Applied Sciences, 1–5.
3. Shamsudin, H., Yusof, U. K., Jayalakshmi, A., & Khalid, M. N. A. (2020). Combining oversampling and undersampling techniques for imbalanced classification. Proceedings of the IEEE International Conference on Control & Automation, 803–808.
4. Feng, F., Li, K.-C., Shen, J., Zhou, Q., & Yang, X. (2020). Using cost-sensitive learning and feature selection algorithms to improve the performance of imbalanced classification. IEEE Access, 8, 69979–69996.
5. Liu, Z., et al. (2020). Self-paced ensemble for highly imbalanced massive data classification. Proceedings of the IEEE International Conference on Data Engineering, 841–852.
6. Park, Y., & Lee, J.-S. (2021). A learning objective controllable sphere-based method for balanced and imbalanced data classification. IEEE Access, 9, 158010–158026.
7. Wang, X., et al. (2021). Evolutionary algorithm-based and network architecture search-enabled multiobjective traffic classification. IEEE Access, 9, 52310–52325.
8. Liu, W., Zhu, C., & Liu, Q. (2022). Multiclass imbalanced and concept drift network traffic classification framework based on online active learning. Engineering Applications of Artificial Intelligence, 117, 105607.
9. Qin, J., Liu, G., & Duan, K. (2022). A new imbalanced encrypted traffic classification model based on CBAM and re-weighted loss function. Applied Sciences, 12(19), 9631.
10. Gu, Y., Yang, Y., & Gao, M. (2023). Learning-based intrusion detection for high-dimensional imbalanced traffic. Computer Communications, 212, 366–376.
11. Rustam, F., & Jurcut, A. D. (2023). Malicious traffic detection in multi-environment networks using novel S-DATE and PSO-D-SEM approaches. Computers & Security, 136, 103564.
12. Ullah, F., Ullah, S., & Lin, J. C.-W. (2023). IDS-INT: Intrusion detection system using transformer-based transfer learning for imbalanced network traffic. Digital Communications and Networks, 9, 287–295.
13. Xu, H., Hu, Y., Cao, W., & Han, L. (2023). An improved jump spider optimization for network traffic identification feature selection. Computers, Materials & Continua, 76(3), 3239–3255.
14. Eldhai, A. M., et al. (2024). Improved feature selection and stream traffic classification based on machine learning in software-defined networks. IEEE Access, 12, 34141–34159.

15. Kang, Q., Chen, X., Li, S., & Zhou, M. (2017). A noise-filtered under-sampling scheme for imbalanced classification. IEEE Transactions on Cybernetics, 47(12), 4263–4274.
16. Gameng, H. A., Gerardo, B. D., & Medina, R. P. (2019). Modified adaptive synthetic SMOTE to improve classification performance in imbalanced datasets. In Proceedings of the IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS) (pp. 1–5).
17. Feng, F., Li, K.-C., Shen, J., Zhou, Q., & Yang, X. (2020). Using cost-sensitive learning and feature selection algorithms to improve the performance of imbalanced classification. IEEE Access, 8, 69979–69996.
18. Liu, Z., Wu, X., Chen, X., & Li, S. (2020). Self-paced ensemble for highly imbalanced massive data classification. In Proceedings of the IEEE 36th International Conference on Data Engineering (ICDE) (pp. 841–852).
19. Shamsudin, H., Yusof, U. K., Jayalakshmi, A., & Khalid, M. N. A. (2020). Combining oversampling and undersampling techniques for imbalanced classification: A comparative study. In Proceedings of the IEEE 16th International Conference on Control & Automation (ICCA) (pp. 803–808).
20. Park, Y., & Lee, J.-S. (2021). A learning objective controllable sphere-based method for balanced and imbalanced data classification. IEEE Access, 9, 158010–158026.
21. Wang, X., Zhang, Y., Li, H., & Zhou, M. (2021). Evolutionary algorithm-based and network architecture search-enabled multiobjective traffic classification. IEEE Access, 9, 52310–52325.
22. Liu, W., Zhu, C., & Liu, Q. (2022). Multiclass imbalanced and concept drift network traffic classification framework based on online active learning. Engineering Applications of Artificial Intelligence, 117, 105607.
23. Qin, J., Liu, G., & Duan, K. (2022). A new imbalanced encrypted traffic classification model based on CBAM and re-weighted loss function. Applied Sciences, 12(19), 9631.
24. Gu, Y., Yang, Y., & Gao, M. (2023). Learning-based intrusion detection for high-dimensional imbalanced traffic. Computer Communications, 212, 366–376.
25. Rustam, F., & Jurcut, A. D. (2023). Malicious traffic detection in multi-environment networks using novel S-DATE and PSO-D-SEM approaches. Computers & Security, 136, 103564.
26. Ullah, F., Ullah, S., & Lin, J. C.-W. (2023). IDS-INT: Intrusion detection system using transformer-based transfer learning for imbalanced network traffic. Digital Communications and Networks, 9, 287–295.
27. Xu, H., Hu, Y., Cao, W., & Han, L. (2023). An improved jump spider optimization for network traffic identification feature selection. Computers, Materials & Continua, 76(3), 3239–3255.
28. Kaur, V., & Virk, A. K. (2024). Hybrid optimization algorithm with random forest for class balancing in network traffic classification. Journal of Electrical Systems, 20(3), 7499–7509.
29. Eldhai, A. M., et al. (2024). Improved feature selection and stream traffic classification based on machine learning in software-defined networks. IEEE Access, 12, 34141–34159.