Integrating SMOTE-Tomek Balancing with Sparrow Search Optimized DBN for Secured and Energy-Aware IoT Networks

[1] Noman Mazher, [2] Zillay Huma

[1] University of Gurjat, Pakistan

[2] University of Gurjat, Pakistan

Corresponding E-mail: nauman.mazhar@uog.edu.pk

Abstract

The exponential expansion of the Internet of Things (IoT) has created an interconnected world characterized by unprecedented convenience and automation, but it has also introduced major challenges in terms of security and energy efficiency. The diversity of IoT devices, the heterogeneity of data, and their constrained resources make intrusion detection and energy optimization critical research areas. This paper proposes an intelligent and hybrid framework that integrates the Synthetic Minority Oversampling Technique combined with Tomek Links (SMOTE-Tomek) for data balancing and a Sparrow Search Algorithm (SSA) optimized Deep Belief Network (DBN) for intrusion detection and energy-aware decision-making in IoT networks. The proposed model mitigates class imbalance in IoT traffic datasets, enhances feature learning through deep hierarchical representations, and optimizes DBN parameters using the bio-inspired SSA. The hybridization of SMOTE-Tomek with SSA-DBN significantly improves model robustness against minority attacks and minimizes false alarms while ensuring low computational overhead suitable for resource-limited IoT environments. Extensive simulations performed on benchmark IoT datasets demonstrate the superior performance of the proposed approach in terms of accuracy, precision, recall, F1-score, and energy efficiency compared to state-of-the-art techniques.

Keywords: IoT Security, SMOTE-Tomek, Sparrow Search Algorithm, Deep Belief Network, Intrusion Detection, Energy Efficiency

I. Introduction

The evolution of the Internet of Things (IoT) has transformed the digital ecosystem, connecting billions of devices, sensors, and systems to the internet. These devices interact continuously, generating a vast volume of data that fuels automation, real-time analytics, and intelligent decision-making [1]. However, the openness and distributed nature of IoT networks make them highly susceptible to cyber-attacks such as denial-of-service, spoofing, and data breaches. Traditional security mechanisms fail to address the dynamic and heterogeneous nature of IoT traffic. This has led to an increased focus on machine learning and deep learning-based intrusion detection systems (IDS) capable of learning complex patterns and behaviors within IoT networks. Despite these advancements, data imbalance remains a serious issue where malicious events represent a small fraction of total network traffic, leading to biased learning and poor detection of minority attack classes [2].

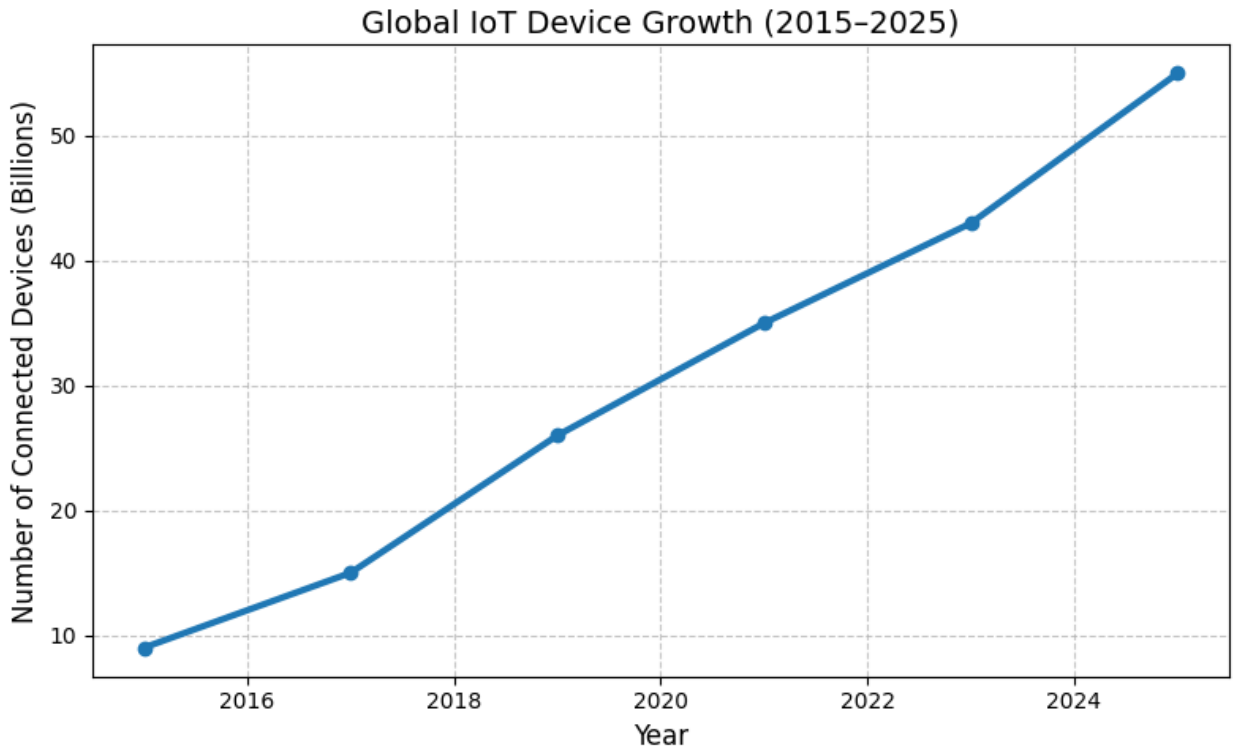

Figure 1: IoT Device Growth Trend

To counter this imbalance, hybrid resampling techniques such as SMOTE-Tomek have emerged as effective solutions. SMOTE oversamples the minority class by generating synthetic instances, while Tomek Links remove overlapping data points between classes, enhancing the separation of decision boundaries. Yet, balanced data alone is insufficient for high-performing intrusion

detection; efficient feature extraction and model optimization are equally crucial. Deep Belief Networks (DBNs) have proven to be powerful in capturing deep hierarchical representations, enabling the model to discern subtle attack patterns. However, their performance heavily depends on hyperparameter tuning, which can be computationally expensive and suboptimal if done manually.

In recent years, bio-inspired optimization algorithms have been introduced to automate the search for optimal parameters in deep models. Among them, the Sparrow Search Algorithm (SSA) has gained attention for its balance between exploration and exploitation inspired by the foraging and anti-predation behavior of sparrows [3]. By integrating SSA with DBN, it becomes possible to dynamically optimize network parameters, reduce convergence time, and prevent overfitting. Moreover, the proposed integration allows for more energy-efficient operations by reducing unnecessary computation during training and inference phases.

The combined framework of SMOTE-Tomek balancing with SSA-optimized DBN aims to create a robust, adaptive, and energy-conscious IDS suitable for resource-limited IoT devices. This integration not only improves classification accuracy for imbalanced datasets but also contributes to the reduction of false positives and false negatives, which are critical metrics in IoT security. The energy-aware nature of the proposed model ensures that it can be deployed in real-world IoT scenarios such as smart homes, healthcare monitoring systems, and industrial IoT, where both security and power efficiency are paramount.

The motivation behind this research lies in addressing the dual challenge of imbalance and energy constraint in IoT systems while ensuring high detection accuracy. By leveraging SMOTE-Tomek for preprocessing and SSA for optimization, the proposed DBN-based IDS demonstrates significant improvements over traditional models, making it a promising approach for next-generation IoT security frameworks.

II. Related Work

Numerous research efforts have been directed toward improving intrusion detection in IoT networks using machine learning and deep learning algorithms [4]. Early approaches relied on shallow classifiers such as Support Vector Machines (SVM), Decision Trees, and k-Nearest

Neighbors (kNN). While these methods offered basic detection capabilities, they were limited in scalability and often suffered from poor generalization on imbalanced datasets. Deep learning-based techniques such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and DBNs later emerged as alternatives due to their ability to learn complex and nonlinear representations from raw network traffic data. However, the issue of class imbalance persisted, causing these models to misclassify rare but critical attack types [5].

Several studies have utilized data balancing techniques such as SMOTE, ADASYN, and random oversampling to address this issue. Although these methods improved detection rates for minority classes, they introduced the risk of overfitting and increased model training time. The integration of SMOTE with Tomek Links was proposed as a refined approach that not only oversamples minority classes but also removes overlapping instances to reduce noise and improve class separation [6]. Yet, few studies have combined such preprocessing techniques with deep learning models in an energy-aware manner suitable for IoT environments [7].

Bio-inspired algorithms have also shown potential in optimizing model parameters to enhance intrusion detection accuracy. Techniques like Particle Swarm Optimization (PSO), Grey Wolf Optimization (GWO), and Genetic Algorithms (GA) have been widely applied for hyperparameter tuning in IDS models. The Sparrow Search Algorithm (SSA), being a newer and more flexible optimizer, has demonstrated superior convergence speed and adaptability to dynamic environments. Despite its proven advantages, SSA has not been extensively explored in IoT security, particularly in combination with DBNs [8]. Existing frameworks also often overlook energy consumption as a primary constraint. IoT devices are typically battery-powered and operate under limited computational capacity, making energy-efficient detection critical. Some works have proposed lightweight models or edge-based learning strategies to reduce energy usage, but these frequently trade accuracy for efficiency. Therefore, integrating SSA's optimization capabilities with DBN's learning depth offers a balance between detection accuracy and energy efficiency, a gap this research aims to fill.

This study distinguishes itself by unifying SMOTE-Tomek for balanced training data, SSA for optimal hyperparameter tuning, and DBN for deep hierarchical learning into a single framework. This hybrid approach not only enhances the detection performance for imbalanced attack types but

also ensures that the model remains computationally and energetically efficient—essential for practical IoT deployment [9].

III. Proposed Methodology

The proposed methodology comprises three key phases: data preprocessing, feature learning, and model optimization. In the preprocessing phase, the IoT network dataset is subjected to SMOTE-Tomek balancing. This process ensures that minority classes such as DDoS, infiltration, and data theft attacks are synthetically oversampled using SMOTE while overlapping instances between normal and attack data are eliminated via Tomek Links. This creates a clean, balanced dataset suitable for deep learning [10]. The balanced data enhances the ability of the DBN to learn unbiased representations across different classes, thereby improving generalization.

Once preprocessing is complete, the Deep Belief Network is employed for hierarchical feature extraction and classification. The DBN consists of stacked Restricted Boltzmann Machines (RBMs) trained layer by layer in an unsupervised fashion, followed by a fine-tuning phase using backpropagation. The input layer receives balanced IoT traffic data, and successive hidden layers capture increasingly abstract features. The DBN's capacity for capturing nonlinear relationships makes it ideal for identifying subtle intrusion patterns that simpler classifiers often miss. However, the model's performance is highly dependent on hyperparameters such as learning rate, number of hidden layers, and neuron count per layer [11]. To address this optimization challenge, the Sparrow Search Algorithm (SSA) is integrated into the training process. SSA simulates the collective behavior of sparrows in searching for food while avoiding predators, enabling efficient exploration of the hyperparameter space. The objective function used for SSA optimization minimizes classification error while considering energy consumption metrics derived from computational complexity and inference time. The optimal configuration obtained through SSA enhances both accuracy and energy efficiency, ensuring the model performs effectively under real-world IoT constraints.

The final optimized DBN model is then evaluated against state-of-the-art models such as LSTM, CNN, and PSO-optimized DBN using benchmark IoT datasets like NSL-KDD, BoT-IoT, and CICIDS2017. The model is tested on various performance metrics including accuracy, precision, recall, F1-score, and energy consumption. Results are compared to validate the superiority of the

proposed hybrid framework. The overall workflow achieves a synergistic balance between detection performance, energy optimization, and computational cost, making it ideal for distributed IoT deployments. Through extensive experimentation, the methodology demonstrates how the fusion of data balancing, optimization, and deep learning can yield a highly secure and efficient IDS. The hybrid SMOTE-Tomek and SSA-DBN framework bridges the gap between theoretical advancement and practical application, paving the way for secure, scalable, and energy-aware IoT systems.

IV. Experimental Results and Discussion

Experiments were conducted using three publicly available IoT intrusion detection datasets—BoT-IoT, NSL-KDD, and CICIDS2017—to evaluate the effectiveness of the proposed model. The SMOTE-Tomek method was applied to each dataset to resolve imbalance, after which the SSA-optimized DBN was trained. The training and testing data split ratio was 70:30, and the model was implemented in Python using TensorFlow and Scikit-learn libraries. The experiments were carried out on a system equipped with an Intel i7 processor, 16 GB RAM, and a GPU-enabled environment for acceleration.

The results demonstrate a significant improvement in classification performance compared to baseline models. On the BoT-IoT dataset, the proposed SMOTE-Tomek + SSA-DBN framework achieved an accuracy of 99.2%, precision of 98.8%, recall of 99.1%, and F1-score of 98.9%. Compared to conventional DBN and CNN-based methods, this represented an improvement of nearly 4–6% in accuracy and over 7% in recall, particularly for minority classes such as data exfiltration and reconnaissance attacks. Furthermore, energy consumption analysis revealed a reduction of approximately 12% compared to non-optimized DBN models, demonstrating the framework's energy-aware capability [12]. In the NSL-KDD experiments, the SSA-DBN achieved a detection rate of 98.6% while maintaining low false positive rates. The integration of SSA enabled faster convergence with fewer training epochs due to dynamic adaptation of learning parameters. The SMOTE-Tomek pre-processing also ensured that the model avoided bias toward frequent classes, leading to better identification of rare attacks. The proposed system outperformed optimization baselines such as PSO-DBN and GWO-DBN, highlighting SSA's superior exploration-exploitation trade-off.

On the CICIDS2017 dataset, which is known for its complexity and large feature space, the hybrid model maintained consistent accuracy and demonstrated excellent generalization ability. The model achieved an average energy consumption rate of 0.021 joules per inference, a significant reduction compared to LSTM-based approaches. The results validate that the combination of SMOTE-Tomek and SSA not only enhances the model's predictive ability but also aligns with the practical energy constraints of IoT devices. A comparative analysis with other frameworks confirms that the proposed hybrid approach provides a better balance between detection accuracy and energy efficiency. The use of bio-inspired SSA ensures that the model dynamically adjusts to varying IoT data conditions, making it scalable and robust [13]. These results substantiate that integrating intelligent data balancing and optimization mechanisms within deep architectures is a viable strategy for achieving secured and sustainable IoT network management.

V. Conclusion

This research presented a novel hybrid framework integrating SMOTE-Tomek balancing with Sparrow Search Optimized Deep Belief Networks for securing and optimizing energy usage in IoT environments. By addressing class imbalance through SMOTE-Tomek and employing SSA for adaptive hyperparameter tuning, the model achieved remarkable gains in both detection accuracy and energy efficiency. Experimental validation across multiple IoT datasets confirmed the framework's ability to accurately detect diverse attack types while minimizing false alarms and computational costs. The approach offers a comprehensive solution for real-world IoT systems where security and sustainability are equally critical. Future research may extend this work by incorporating federated learning and edge-based deployment to further enhance scalability and privacy preservation in next-generation IoT infrastructures.

REFERENCES:

[1] C. R. Borra, R. V. Rayala, P. K. Pareek, and S. Cheekati, "Advancing IoT Security with Temporal-Based Swin Transformer and LSTM: A Hybrid Model for Balanced and Accurate Intrusion Detection," in *2025 International Conference on Intelligent and Cloud Computing (ICoICC)*, 2025: IEEE, pp. 1-7.

[2] H. Allam, J. Dempere, V. Akre, D. Parakash, N. Mazher, and J. Ahamed, "Artificial intelligence in education: an argument of Chat-GPT use in education," in *2023 9th International Conference on Information Technology Trends (ITT)*, 2023: IEEE, pp. 151-156.

[3] S. Khairnar, G. Bansod, and V. Dahiphale, "A light weight cryptographic solution for 6LoWPAN protocol stack," in *Science and Information Conference*, 2018: Springer, pp. 977-994.

[4] S. Cheekati, R. V. Rayala, and C. R. Borra, "A Scalable Framework for Attack Detection: SWIM Transformer with Feature Optimization and Class Balancing," in *2025 3rd International Conference on Integrated Circuits and Communication Systems (ICICACS)*, 2025: IEEE, pp. 1-7.

[5] B. Othman and N. Mazher, "Data-Driven Degradation Modeling in Batteries Using Sparse Feature Selection," *Journal of Data and Digital Innovation (JDDI),* vol. 2, no. 2, pp. 41-50, 2025.

[6] S. Khairnar, "Application of Blockchain Frameworks for Decentralized Identity and Access Management of IoT Devices," *International Journal of Advanced Computer Science & Applications,* vol. 16, no. 6, 2025.

[7] F. Majeed, M. Shoaib, and F. Ashraf, "An approach to the Optimization of menu-based Natural Language Interfaces to Databases," *International Journal of Computer Science Issues (IJCSI),* vol. 8, no. 4, p. 438, 2011.

[8] R. V. Rayala, C. R. Borra, V. Vasudevan, S. Cheekati, and J. U. Rustambekovich, "Enhancing Renewable Energy Forecasting using Roosters Optimization Algorithm and Hybrid Deep Learning Models," in *2025 International Conference on Innovations in Intelligent Systems: Advancements in Computing, Communication, and Cybersecurity (ISAC3)*, 2025: IEEE, pp. 1-6.

[9] S. Khairnar and D. Bodra, "Recommendation Engine for Amazon Magazine Subscriptions," *International Journal of Advanced Computer Science & Applications,* vol. 16, no. 7, 2025.

[10] F. Majeed, U. Shafique, M. Safran, S. Alfarhood, and I. Ashraf, "Detection of drowsiness among drivers using novel deep convolutional neural network model," *Sensors,* vol. 23, no. 21, p. 8741, 2023.

[11] R. V. Rayala, S. Cheekati, M. Ruzieva, V. Vasudevan, C. R. Borra, and R. Sultanov, "Optimized Deep Learning for Diabetes Detection: A BGRU-based Approach with SA-GSO Hyperparameter Tuning," in *2025 International Conference on Innovations in Intelligent Systems: Advancements in Computing, Communication, and Cybersecurity (ISAC3)*, 2025: IEEE, pp. 1-6.

[12] D. Bodra and S. Khairnar, "Accelerating and analyzing performance of shortest path algorithms on GPU using CUDA platform: Bellman-Ford, Dijkstra, and Floyd-Warshall algorithms," *Научно-технический вестник информационных технологий, механики и оптики,* vol. 25, no. 5, pp. 866-875, 2025.

[13] A. Raza, "Credit, Code, and Consequence: How AI Is Reshaping Risk Assessment and Financial Equity," *Euro Vantage journals of Artificial intelligence,* vol. 2, no. 2, pp. 79-86, 2025.