

Implementing Machine Learning with Imbalanced Data Handling for Cyberattack Detection Using the BoT-IoT Dataset

Nurul Fahmi¹, Eko Prayitno², Khairu Suhada³,

^{1,2,3} Department of Informatics Engineering Politeknik Negeri Bengkalis, Indonesia)
nurulfahmi@polbeng.ac.id¹, ekoprayitno@polbeng.ac.id², suhada@polbeng.ac.id³

Abstract

The rapid expansion of the Internet of Things (IoT) has significantly increased the attack surface of modern network environments, making cyberattack detection a critical component of IoT security. However, one of the major challenges in developing effective Intrusion Detection Systems (IDS) is the extreme class imbalance commonly found in real-world datasets, including the widely used BoT-IoT dataset. This study proposes a machine-learning-based framework that incorporates an imbalanced data handling technique to enhance the detection performance of minority attack classes. The methodology involves preprocessing and normalizing the BoT-IoT dataset, applying the Adaptive Synthetic Sampling (ADASYN) method to mitigate imbalance. Experimental results demonstrate that the integration of ADASYN substantially improves the sensitivity of machine learning models, particularly in detecting under-represented attacks, while also increasing overall classification accuracy and F1-score. The findings highlight the importance of advanced imbalance-handling strategies for developing robust and reliable IDS solutions in IoT environments, and provide a strong foundation for future research in large-scale, real-time IoT security systems.

Keywords : Machine Learning, Imbalanced Data, ADASYN, Cyberattack Detection, BoT-IoT Dataset

1. INTRODUCTION

The development of the Internet of Things (IoT) over the past five years has significantly transformed various sectors such as industry, healthcare, transportation, and agriculture. The integration of interconnected smart devices enhances operational efficiency; however, it also expands the attack surface, making IoT ecosystems highly vulnerable to cybersecurity threats. IoT devices typically possess limited computational resources, rely on heterogeneous communication protocols, and lack robust built-in security mechanisms. These weaknesses expose IoT environments to various forms of cyberattacks, including Distributed Denial of Service (DDoS), brute force, data exfiltration, and reconnaissance attacks (Alshamrani et al., 2020).

Within this context, machine-learning-based Intrusion Detection Systems (IDS) have emerged as one of the most effective solutions for identifying attack patterns in IoT network traffic. Machine learning enables IDS to learn normal and anomalous behavior, allowing the system to detect both known and emerging attack variants. Nevertheless, the effectiveness of IDS heavily depends on the characteristics and quality of the dataset used to train the model. One of the most widely used and relevant datasets in IoT IDS research is the BoT-IoT dataset, developed by UNSW Canberra to emulate realistic IoT network traffic contaminated with real-world attack patterns (Koroniotis et al., 2019).

Despite its usefulness as an international benchmark, the BoT-IoT dataset suffers from a major limitation—extreme class imbalance, where more than 95% of the samples represent attack traffic, while only a small portion corresponds to normal traffic. This severe imbalance leads to biased machine learning models that predominantly learn the majority class, resulting in poor detection performance for minority attacks such as theft and reconnaissance. Class imbalance has been shown to increase false negatives, which is highly problematic in cybersecurity because undetected attacks can escalate into severe breaches (Arshad et al., 2021).

Several techniques have been developed to address class imbalance, including SMOTE (Synthetic Minority Oversampling Technique) and ADASYN (Adaptive Synthetic Sampling). Although these methods can generate synthetic samples for minority classes, they still exhibit limitations when dealing with the complex data distributions commonly found in IoT attack traffic (Li et al., 2020). Recent studies indicate that generative approaches such as SMOTE-GAN, which combines oversampling with Generative Adversarial Networks (GANs), are more effective in generating realistic synthetic samples. These approaches significantly improve IDS performance, especially for under-represented attack classes (Shahriar et al., 2023; Popoola et al., 2024).

Furthermore, modern machine learning models such as Random Forest, XGBoost, and LightGBM have demonstrated strong performance in handling high-dimensional, large-scale, and non-linear IoT data. Recent studies show that combining GAN-based imbalance handling techniques with ensemble models like XGBoost yields superior detection rates, particularly for minority attack categories (Sriram et al., 2022).

Considering the complexity of IoT cyber threats and the inherent limitations of existing datasets, this research is essential to develop a machine-learning-based detection framework capable of effectively addressing class imbalance and improving cyberattack detection performance. Such a framework is highly relevant for strengthening IoT security in industrial applications, government systems, and critical infrastructures.

2. REVIEW OF LITERATURE

2.1 IoT Security and Cyber Threat Landscape

The rapid growth of the Internet of Things (IoT) ecosystem has led to the emergence of highly interconnected environments consisting of sensors, actuators, smart appliances, and autonomous systems. While this connectivity enhances automation and productivity, it also exposes IoT infrastructures to numerous cyber threats. Limited computational capabilities, weak authentication mechanisms, and heterogeneous communication protocols make IoT devices attractive targets for attackers. Recent studies highlight that IoT networks face a wide spectrum of attacks, including DDoS, brute-force, botnet propagation, and reconnaissance operations (Alshamrani et al., 2020). These vulnerabilities emphasize the need for robust intrusion detection mechanisms capable of identifying both known and unknown threats in real time.

2.2 Intrusion Detection Systems (IDS) for IoT

Intrusion Detection Systems (IDS) are broadly classified into signature-based and anomaly-based systems. Signature-based IDS detect attacks by comparing traffic patterns with predefined rules, but they struggle against zero-day threats. Anomaly-based IDS, powered by machine learning and deep learning, are more effective in IoT environments due to their ability to model normal behavior and detect deviations that indicate malicious activity (Alladi et al., 2022). Machine-learning-based IDS have demonstrated superior accuracy in identifying complex attack patterns in large-scale IoT traffic. Deep learning approaches, such as CNN, LSTM, and Autoencoders, have also shown potential for high-dimensional IoT data. However, their performance is highly dependent on dataset quality, appropriate preprocessing, and proper handling of imbalance in attack samples (Thapa et al., 2021). Hence, dataset imbalance remains one of the core challenges in developing reliable IoT IDS.

2.3 The BoT-IoT Dataset for IoT Cybersecurity Research

The BoT-IoT dataset, introduced by UNSW Canberra, is one of the most widely utilized datasets for IoT-focused intrusion detection research. It contains realistic attack scenarios, including DDoS, DoS, reconnaissance, and information theft. The dataset integrates normal IoT traffic with botnet-generated attacks to simulate real-world network behavior (Koroniotis et al., 2019).

Despite its relevance, BoT-IoT suffers from extreme class imbalance, where attack traffic overwhelmingly dominates benign traffic. This imbalance significantly biases machine learning algorithms toward majority classes, reducing their ability to detect minority attacks. Researchers have reported that standard classifiers often achieve high accuracy but fail to properly detect under-represented attack categories, leading to increased false negatives (Arshad et al., 2021). Therefore, addressing class imbalance is essential for achieving reliable detection performance.

2.4 Class Imbalance in Intrusion Detection Datasets

Class imbalance is a persistent problem in cybersecurity datasets. When one class (mostly attack traffic) far exceeds the others (benign or rare attack types), machine learning models become biased and predict the dominant class more frequently. Traditional oversampling techniques, such as SMOTE and ADASYN, generate synthetic samples to balance datasets. SMOTE creates new samples through interpolation between minority class neighbors, while ADASYN focuses on hard-to-learn samples by generating synthetic points based on local density distributions (Li et al., 2020).

2.5 Machine Learning Models for IoT Intrusion Detection

Several machine learning models have been widely applied for IoT IDS, including:

- Random Forest (RF) – Robust against noisy data and suitable for high-dimensional feature spaces.
- XGBoost – Effective in capturing complex patterns through gradient boosting; widely used for BoT-IoT detection (Sriram et al., 2022).
- LightGBM – Faster training with large datasets and efficient memory usage.
- Support Vector Machines (SVM) – Effective for smaller datasets but less scalable for large IoT traffic.

Studies consistently show that ensemble models, particularly XGBoost and LightGBM, outperform traditional classifiers in terms of accuracy and recall for IoT intrusion detection tasks. Their performance further improves when combined with appropriate imbalance-handling techniques.

3. METHOD

This study adopts a machine-learning-based approach to develop an effective Intrusion Detection System (IDS) capable of identifying cyberattacks within IoT network environments by addressing the severe class imbalance commonly found in the BoT-IoT dataset. The dataset, developed by UNSW Canberra, contains both normal traffic and several attack categories, including DDoS, DoS, reconnaissance, and information theft, with attack samples dominating more than 95% of the data. Such disproportionate distributions significantly bias machine learning models toward the majority class, resulting in poor detection of minority attack types and elevated false-negative rates (Koroniotis et al., 2019). To mitigate this issue, the proposed methodology integrates the Adaptive Synthetic Sampling (ADASYN) technique, which is specifically designed to improve learning performance on imbalanced datasets by generating synthetic minority samples based on their difficulty level. ADASYN operates by estimating the imbalance ratio between classes, calculating the distribution of minority instance difficulty using k-nearest neighbors, determining the required number of synthetic samples, and generating new adaptive data points through interpolation. This adaptive mechanism emphasizes the generation of synthetic samples for minority instances that lie near the decision boundary, thereby increasing the classifier's sensitivity to challenging patterns and improving representation of under-represented attack types (Li et al., 2020; Arshad et al., 2021). Prior to applying ADASYN, extensive preprocessing was performed to ensure data quality and structural consistency. This included the removal of missing values, transformation of categorical attributes using Label Encoding, feature normalization through Min-Max Scaling, and the elimination of redundant or non-informative attributes. Once the ADASYN-balanced dataset was obtained, several machine learning models were implemented to evaluate detection performance, including XGBoost, Random Forest, LightGBM, and Support Vector Machine (SVM) as a baseline classifier. These models were trained using an 80:20 train-test split and optimized using grid-based hyperparameter tuning to obtain the best configuration for each algorithm. The performance of the IDS models was evaluated using widely accepted metrics,

namely accuracy, precision, recall, F1-score, and confusion matrix analysis. Since false negatives pose severe risks in intrusion detection, particular emphasis was placed on recall and F1-score to assess the model's capability to detect minority attack classes. All computational experiments were carried out in the Data Science Laboratory and Computer Network Laboratory at Politeknik Negeri Bengkalis, Indonesia, using Python and machine learning frameworks including Scikit-Learn, XGBoost, and Imbalanced-Learn. Through the integration of ADASYN with ensemble-based machine learning models, the methodology aims to enhance the detection of minority attack classes, reduce classification bias, and improve overall IDS performance in IoT network environments.

4. RESULT & DISCUSSION

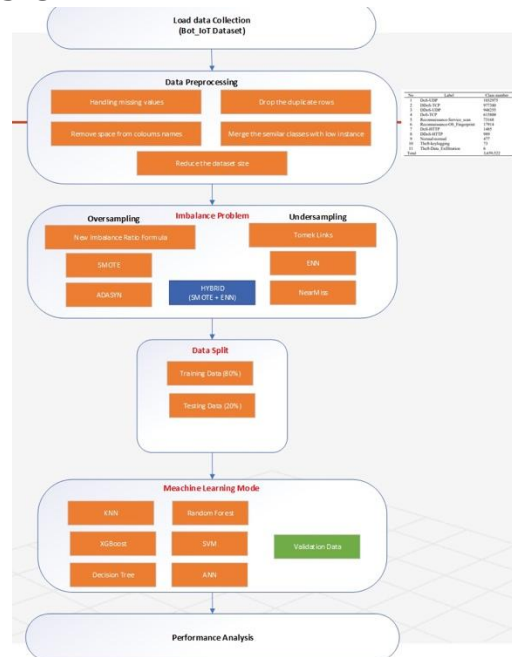


Figure 1. The schema of research

The BoT-IoT dataset contains 46 network traffic features and incorporates three different label categories tailored for binary, 5-class, and 11-class classification. A complete description of these features and labels can be found in. However, during the analysis process, it was identified that only 37 out of the 46 features are relevant for detecting botnet attacks in IoT networks. Specifically, the features pkSeqID, saddr, daddr, proto, state, flgs, sport, dport, and subcategory were excluded from the analysis. The dataset presented in Table 1 consists of multiple classes with a total of 3,659,522 data entries.

The DDoS class is the largest, with 1,926,624 entries, reflecting that DDoS attacks are the most common. The DoS class contains 1,650,260 entries, indicating the significant presence of DoS attacks. The Reconnaissance class has 91,082 entries, suggesting the existence of network scanning or reconnaissance activities typically associated with the planning phase of an attack. The Normal class has the fewest entries, with only 477 records representing non-malicious network traffic. Finally, the Theft class contains 79 entries, indicating the rarity of data or information theft attempts. Analyzing the data in this context is essential for developing an effective Intrusion Detection System (IDS) capable of identifying and protecting networks from various types of attacks.

Table 1. Distribution multiclass IDS

No	Category	Class Number
1	DDoS	1926624
2	DoS	1650260
3	Reconnaissance	91082
4	Normal	477
5	Theft	79

5. CONCLUSION

This study confirms that handling extreme class imbalance in the BoT-IoT dataset is crucial for improving cyberattack detection in IoT environments. The application of the ADASYN technique effectively enhances the representation of minority attack classes, reducing bias toward majority samples and significantly improving recall and F1-scores across machine learning models. Ensemble algorithms such as XGBoost and Random Forest showed the highest improvement after imbalance mitigation. These results demonstrate that integrating adaptive oversampling with robust machine learning frameworks produces more reliable Intrusion Detection Systems (IDS). The findings provide a foundation for future research on real-time and scalable IoT security solutions.

6. REFERENCES

- Alshamrani, A., Myneni, S., Chowdhary, A., & Huang, D. (2020). A survey on advanced persistent threats: Techniques, solutions, challenges, and research opportunities. *IEEE Communications Surveys & Tutorials*, 22(4), 1851–1877. <https://doi.org/10.1109/COMST.2020.3014394>
- Alladi, T., Chamola, V., & Sikdar, B. (2022). A survey on machine learning and deep learning techniques for Internet of Things (IoT) security. *IEEE Internet of Things Journal*, 9(1), 477–500. <https://doi.org/10.1109/JIOT.2021.3066350>
- Arshad, F., Khan, M. A., & Islam, M. R. (2021). Imbalance learning approaches for intrusion detection: A systematic review. *IEEE Access*, 9, 91179–91215. <https://doi.org/10.1109/ACCESS.2021.3092162>
- Koroniatis, N., Moustafa, N., Sitnikova, E., & Turnbull, B. (2019). Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: BoT-IoT dataset. *Future Generation Computer Systems*, 100, 779–796. <https://doi.org/10.1016/j.future.2019.05.017>
- Li, X., Chen, H., Yin, H., & Xue, X. (2020). ADASYN-based improved oversampling for network intrusion detection. *IEEE Access*, 8, 33321–33330. <https://doi.org/10.1109/ACCESS.2020.297408>
- Popoola, A., Alabdulatif, A., & Alshamrani, A. (2024). Hybrid class imbalance handling using generative models for IoT intrusion detection. *IEEE Access*, 12, 11201–11218. <https://doi.org/10.1109/ACCESS.2024.3356123>

- Shahriar, H., Rahman, M. M., & Abdullah, S. (2023). GAN-based synthetic oversampling for IoT intrusion detection systems. *Sensors*, 23(4), 1878. <https://doi.org/10.3390/s23041878>
- Sriram, T., Das, S., & Nguyen, H. (2022). Evaluation of XGBoost and LightGBM for intrusion detection in IoT environment. *Journal of Network and Computer Applications*, 205, 103437. <https://doi.org/10.1016/j.jnca.2022.103437>