Journal of Robotics and Control (JRC)
Volume 3, Issue 4, July 2022
ISSN: 2715-5072, DOI: 10.18196/jrc.v3i4.

A Robust Ensemble Learning for DDoS Attack Classification on the Internet of Medical Things

Nisrina Nurhaliza [1], Sudianto Sudianto [2,*], Pradana Ananda Raharja [3]
[1, 2, 3] Department of Informatics, Telkom University, Purwokerto, Indonesia
*Corresponding Author: sudianto@telkomuniversity.ac.id

***Abstract*—Distributed Denial of Service (DDoS) attacks pose significant threats to the Internet of Medical Things (IoMT), potentially disrupting critical healthcare services. A significant challenge in detecting these attacks is the high imbalance in network traffic data, which can bias classification models. This study introduces a hybrid approach that integrates the XGBoost algorithm with Principal Component Analysis (PCA) and undersampling to address class imbalance and enhance detection performance. Three configurations were tested: (1) XGBoost + PCA + Undersampling, (2) XGBoost + PCA, and (3) XGBoost + Undersampling. The model was evaluated using a publicly available multiclass DDoS dataset under an 80:20 training-testing split. The XGBoost + Undersampling method achieved the highest performance, with accuracy, precision, recall, and F1-score of 99.98%. Despite these results, potential limitations—such as data loss due to undersampling and excluding cross-validation or external testing. Thus, the proposed ensemble learning technique has proven to robustly improve the performance of DDoS attack detection in unbalanced dataset conditions.**

Keywords—DDoS detection; ensemble learning; class imbalance; IoMT security; XGBoost

I. INTRODUCTION

Internet of Things (IoT)-based devices are seeing a 19% increase in connected devices yearly. By 2027, around 47% of IoT applications will integrate artificial intelligence (AI) elements as innovative solutions in problem-solving through integrated technology. As part of the internet ecosystem, IoT has contributed to more accurate and structured data management. In addition, the automation of IoT devices enables increased efficiency and productivity in various aspects of life and drives the transformation of digital services [1], [2]. However, along with the rapid growth of IoT devices, security challenges also arise, including the threat of cyber-attacks that can originate from IoT devices [3]. These attacks can occur intentionally or unintentionally by overloading resources and networks [4], [5]. Distributed Denial of Service (DDoS) attacks target IoT vulnerabilities [6]. Distributed Denial of Service (DDoS) is an attack that aims to disrupt internet service availability so that authorized users cannot access it [4]. Attackers flood the target with internet traffic, making it difficult for the server to handle legitimate requests due to excessive resource consumption [7]. As technology evolves, DDoS attacks become increasingly complex by utilizing infected IoT devices as part of a botnet, allowing attacks to be carried out widely and coordinated [8]. Detection of DDoS attacks is often difficult unless the attacker uses the same IP address repeatedly in multiple access attempts. The main challenge in overcoming DDoS attacks arises from their association with high-security services and outdated attack protocols, which can overload the network backbone [9].

In the case of the Internet of Medical Things (IoMT), as part of the IoT ecosystem in healthcare, it is vulnerable to Distributed Denial of Service (DDoS) attacks. IoMT integrates medical devices, Wireless Body Area Networks (WBAN), artificial intelligence (AI), and cloud technology to monitor health conditions in real-time [10].

IoMT has much more complex security challenges including critical medical devices such as pacemakers, insulated pumps and patient health monitoring systems that have a low tolerance to system interference and failure [11], the vulnerability of IoMT devices to DDoS attacks is different from attacks on common IoT systems where IoMT has a direct impact on the risks associated with Generative AI and 5G-IoT which if exploited can lead to data breaches, unauthorized access and errors in commands, controls, and potential device hazards that can endanger patient safety, while in general IoT, it usually has an impact on non-critical data or services [12]. Key components of IoMT, such as data collection, storage, transfer, and analysis, play a role in early warning systems to detect and control the spread of diseases [13]. Sensors on end-user devices, such as mobile phones, tags, or health monitors, send data to the cloud for analysis and medical decision-making [14]. However, high connectivity between devices with diverse communication methods and protocols significantly increases the volume of network traffic. The large-scale data generated can potentially cause data imbalances, especially in real-time networks with high data dimensions. This imbalance is increasingly complex due to the global interactions between devices in the IoMT ecosystem [15]. This condition increases the risk of DDoS attacks, which use high traffic to disrupt services and overload network infrastructure. Therefore, early mitigation is needed to discover the attacks that occur using an artificial intelligence approach. Most DDoS detection techniques developed in IoT environments have not been fully adapted to the specific needs of IoMT (Figure 1). Conventional methods tend to ignore the limitations of computing power, energy consumption, and the need for real-time response that are typical characteristics of IoT-based medical systems [16]. In addition, conventional machine learning methods are often prone to overfitting and require large amounts of training data, which are not always available

in a medical context due to limited data access and privacy [17].

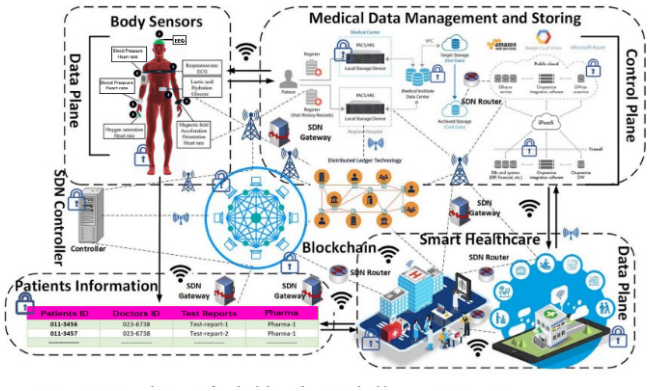

Fig. 1. Architecture Internet of Medical Things

One of the methods of artificial intelligence is Ensemble learning. Ensemble Learning is an optimization technique in Machine Learning that combines several classification models to improve prediction accuracy [18], [19]. The ensemble learning method has shown superior performance in many studies, but it is often overlooked that this approach has inherent drawbacks such as high model complexity, overfitting risk, and large computing requirements—all of which are serious obstacles to IoMT-based edge devices [20]. This method has various approaches, such as Bagging, Boosting, Stacking, AdaBoost, and XGBoost, as well as the latest type that is dynamic [21]. In the context of DDoS attacks on IoMT devices, the Ensemble Learning method addresses dataset imbalances by developing new techniques, including data augmentation and feature optimization. Dataset imbalances in the classification of DDoS attacks can cause the model to be more likely to recognize the majority class, thus ignoring the minority class. This approach aims to improve the quality of predictions and produce a more optimal classification model [22]. Therefore, there is still a significant gap in the development of a DDoS detection approach that is not only accurate, but also efficient and contextual against resource-constrained IoMT environments by proposing an ensemble learning approach reinforced with Principal Component Analysis (PCA) techniques and tailored undersampling strategies. This strategy aims to reduce data complexity and address multiclass imbalances in DDoS attack datasets, while maintaining detection performance at a high level and efficient with real-time applicability capabilities due to XGBoost's lightweight inference [23].

Data imbalances where the amount of attack data is much less than normal data becomes critical because the potential failure to detect even a single attack can have a fatal impact in the context of IoMT [24]. In addition, high-dimensional network traffic deteriorates detection performance, requiring dimension reduction techniques such as Principal Component Analysis (PCA) to filter out the most relevant features without sacrificing detection accuracy. However, the application of undersampling in the context of healthcare poses ethical and practical dilemmas. Reducing data from the majority class can lead to the loss of important information that can accurately represent normal traffic patterns [25]. This can degrade the reliability of the detection system, which is highly undesirable in a medical environment that demands high precision. Therefore, undersampling strategies must be carefully designed to maintain information integrity while addressing data imbalances. Based on these challenges, this study aims to propose a robust classification approach to detect DDoS attacks on IoMT data, by optimally integrating PCA and undersampling techniques.

II. Related Work

Distributed Denial-of-Service (DDoS) attacks cause significant impacts on various sectors, including economic losses, network disruptions, and damage to server infrastructure that hackers use to launch cyberattacks. In recent years, research has proposed various methods of detecting DDoS attacks to minimize their impact. This detection approach can be categorized based on the application environment, namely conventional networks, cloud environments, and Software-Defined Networking (SDN) [26], [27], [28]. SDN is a modern network architecture that allows the management of data traffic through OpenFlow as a communication protocol to regulate data exchange between routers and switches through controllers [29]. Previous studies proposed a DDoS attack mitigation approach by optimizing multi-controller SDN-based machine learning, using K-Means++ and OPTICS algorithms in real network topologies [30].

DDoS attacks can hinder the correlation of data traffic, especially on datasets that experience class imbalances. Several studies have developed hybrid sampling solutions by combining intelligent undersampling and oversampling with the SMOTE method, demonstrating improved efficiency in detecting DDoS attacks on large datasets [31]. In addition, other studies have shown that oversampling techniques in analog-to-digital converter systems effectively increase the sum rate by reducing the impact of quantization distortion in hardware [32]. The implementation of faster and more scalable detection methods is needed because the traditional approach of using a single controller is still less effective in dealing with increasingly sophisticated cyberattacks. Therefore, another study has evaluated Mininet with POX Controller in a simulation of a real-time setup-based DDoS attack detection environment with evaluation matrix analysis to assess system performance [33].

Researchers have developed a cloud computing-based detection approach to improve the effectiveness of DDoS attack identification. Previous studies have proposed an ensemble learning-based cyberattack detection framework with fog-cloud architecture, which combines Decision Tree, Naïve Bayes, and Random Forest as base classifiers, as well as XGBoost as meta-classifiers [34], [35], [36]. This model was tested using the ToN-IoT dataset and achieved an accuracy of 96.35% with a detection rate of 99.98% [37] [38]. However, this study focuses more on detecting cyberattacks without highlighting the specific DDoS threat on the Internet of Medical Things (IoMT) [39]. Another study proposes a similar approach for IoMT with an accuracy of 94.74%. However, the feature selection strategy is still limited, so the model's efficiency in handling complex and dynamic data in IoMT environments can still be improved [40]. In addition,