CHAPTER I
INTRODUCTION

1. 1 Background

SDN (Software Defined Network) is a new approach in network programming in designing, building, and managing computer networks by separating the control plane and data plane[1]. Instead of traditional networks, SDN separates the control plane from the data plane in order to centralize the network with the aim of concentrating all settings on the control plane. SDN has the ability to dynamically control, modify, and manage network behavior through open interface software [2]. SDN was developed during a time when traditional networks were particularly intricate, whereas contemporary networks that sustain application development require adaptable and simple monitoring capabilities. SDN empowers users to oversee and sustain unified networks, assembling and integrating network devices within a concise period of time. SDN can also expedite the swift formulation, preparation, and administration of devices and measures at a reduced cost.

Network development has frequently been impeded by cybersecurity concerns and cyber threats. Unlike conventional networks, SDN possesses various benefits. One of these includes the ability to implement a centralized control function whereby a switch receives all packets originating from the host, the packet delivery path of which has been set by the controller, placing the controller at the core of SDN's operation. However, due to its centralized nature, SDN is vulnerable to attacks, including Distributed Denial of Service (DDoS) attacks [3]**.** DDoS attacks pose a considerable danger to computer networks. Cyber-attacks aimed at disrupting network availability and performance include DDoS attacks which target network resources and render certain services inaccessible, leading to network degradation [4]. DDoS attacks have a considerable impact on the uptime of SDN systems. Specifically, the SDN controller experiences the greatest impact since it represents the most vulnerable point of failure [5].

Recently, there has been significant research on identifying DDoS attacks on SDN networks using machine learning techniques. Several studies explored this area, seeking to improve the accuracy and effectiveness of detection methods. In 2019 [6], the researcher undertook an experiment to detect DDoS attacks by proposing several machine-learning algorithms. These algorithms included Support Vector Machine (SVM) and Random Forest

(RF). The results revealed that the employed machine learning algorithms attain a high degree of detection accuracy, albeit not perfect. Elsayed et al. [7] The research employed a machine-learning approach to identify DDoS attacks on SDN. Several machine learning algorithms were analyzed, including SVM, J48, Naïve Bayes (NB), and RF. The results indicate that the J48 algorithm has a significantly higher detection accuracy when compared to other algorithms like RF. Rahman et al.[8] also implemented the J48 algorithm to identify and prevent DDoS attacks on SDN. In addition to J48, various machine learning algorithms such as RF, SVM, and K-Nearest Neighbors (KNN) were utilized by the authors. The obtained results revealed that the J48 algorithm provides the highest detection accuracy and outperforms the pre-defined machine learning algorithms in detecting and blocking DDoS attacks on SDN.

Dias Firdaus et al. [3] Discuss how to detect DDoS attacks on SDN using ensemble methods in machine learning algorithms, specifically K-Means +++ and RF to increase accuracy results and proven by the results obtained are very satisfying. The ensemble method of machine learning algorithms used in this research obtained a perfect detection accuracy of 100%. The authors recommend incorporating the RF algorithm with hyper-tuning in future research for detecting DDoS attacks on SDN. In addition, Hassan A. Alamri and Vijey Thayananthan [9] performed a review and analysis of machine learning schemes to secure SDN environments from DDoS attacks. Using the dataset from CIC-DDoS 2019, the authors analyzed machine learning using several algorithms namely NB, K-NN, SVM, Decision Tree (DT), RF, and Extreme Gradient Boosting (XGBoost). This research obtained the highest accuracy results in the XGBoost algorithm which was then followed by other algorithms namely RF, NB, KNN, SVM, and DT. Altamemi AJ et al. [5] also evaluate the research that has been done before and then conduct research using machine learning to quickly detect DDoS attacks. The dataset used in this research is a real-time dataset in the form of CSV. The machine learning algorithms used to detect DDoS attacks are DT, Linear Regression (LR), and NB algorithms and it is concluded that the algorithm is more effective for detecting DDoS attacks with an accuracy rate of 99.90%.

Due to the potential severity of the consequences, a DDoS attack on an SDN network can have catastrophic effects [10], a solution is needed to improve the security of SDN networks. This research aims to accomplish DDoS attack detection by utilizing machine learning algorithms on SDN networks. This research aims to evaluate the performance of Random Forest in DDoS attack detection experiments using real datasets. Additionally, this

research focuses on improving the detection accuracy value of the Random Forest algorithm and optimizing the performance of the Random Forest algorithm using the Hyperparameter Tuning method by considering various performance metrics. By adopting a comprehensive approach, the present research aims to enhance the security of SDN networks. It aims to assist technicians and scholars in detecting DDoS attacks on SDN networks and serve as a point of reference for future research development.

1. 2 Problem Identification

Based on the background presented, this research aims to identify the problem in SDN networks.

1. How to choose influential hyperparameters to optimize the performance of the Random Forest algorithm?
2. How to determine the best hyperparameter value to optimize the performance of the Random Forest algorithm and obtain a high accuracy value of DDoS attack detection?

1. 3 Objective and Contribution

Referring to the problem identification that has been described, the objectives of this research are as follows:

1. To determine useful hyperparameters to optimize the performance of Random Forest.
2. To determine the optimal value of each hyperparameter using Hyperparameter Tuning Random Search to optimize the performance of Random Forest.
3. To demonstrate that the Hyperparameter Tuning Random Forest method is effective in detecting DDoS attacks.

1. 4 Scope of Work

To limit the scope of the discussion, these issues are defined as follows:

1. The Random Forest algorithm is tested for its effectiveness in detecting DDoS attacks on SDN networks.
2. The dataset used is InSDN data obtained from the Virtual SDN Testbed Network.
3. Optimising the Random Forest model's performance involves using the Hyperparameter Tuning method.
4. Random Search is employed as a technique for Hyperparameter Tuning to reduce the time required to determine the most optimal Hyperparameter.

1. 5 Hypothesis

In SDN networks, the controller is highly susceptible to attacks. Although several attack detection techniques have been proposed, only a few consider the security of the controller [10]. The Random Forest algorithm was chosen for this research due to its effectiveness in achieving high accuracy and efficiency in processing large amounts of data. Significant hyperparameters for the performance of the Random Forest algorithm will be selected. Hyperparameter Tuning can then be used to optimize the algorithm's performance by finding the best value for the selected hyperparameters. Random Search will be used as a method for Hyperparameter Tuning to determine the optimal value for each hyperparameter. This technique is commonly used in machine learning to create a more optimized and higher quality model, resulting in more accurate detection results.

1. 6 Methodology

The methodology chosen to be used in completing this research is the Design Science Research Process (DSRP) [11]. The steps taken in completing this research are a literature review to understand and identify the problems to be discussed, then determine a solution and the goal of the solution, namely, making a system design from the selected solution, namely making a machine learning-based security system to detect DDoS attacks in an SDN network, then demonstrating the use and results of the research based on the results of the analysis in the form of performance metrics accuracy, precision, recall, training time and F1-Score. Furthermore, the evaluation process is carried out based on the research results obtained to draw a final conclusion.

1. 7 Research Methodology

To achieve high accuracy detection and fast data processing times, a machine learning model is utilized. The Random Forest algorithm was selected as the machine learning algorithm of choice. This decision was made due to its ease of implementation and ability to adapt to large data sets. By utilizing this algorithm, a heightened level of accuracy can be achieved [12]. Hyperparameter tuning is additionally utilized in optimizing the Random Forest performance in machine learning. Furthermore, the authors conducted data processing and testing using the Random Forest algorithm on the InSDN dataset. The dataset has been divided into three distinct groups: the first group comprises datasets with normal traffic, the second

group includes datasets with attack traffic directed at Mealspotable-2, and the final group consists of datasets with attacks on the OVS machine under consideration [13].