A HYBRID MULTI-TIER APPROACH FOR IOT BOTNET DETECTION AND ENHANCED RISK ASSESSMENT

ASHRAF SULIEMAN ALI MASHALEH

UNIVERSITI SAINS MALAYSIA

2025

A HYBRID MULTI-TIER APPROACH FOR IOT BOTNET DETECTION AND ENHANCED RISK ASSESSMENT

by

MASHALEH ASHRAF SULIEMAN ALI

Thesis submitted in fulfillment of the requirements
for the degree of
Doctor of Philosophy

January 2025

ACKNOWLEDGEMENT

{سورة يوسف} {نَرْفَعُ دَرَجَاتٍ مِّن نَّشَاء وَفَوْقَ كُلِّ ذِي عِلْمٍ عَلِيمٌ -76}

First and foremost, I would like to thank Allah, our Creator, for His bounties and for providing me with the fortitude and persistence to complete my Ph.D. His advice and assistance have been invaluable in helping me navigate the hurdles and obstacles I experienced along the way. I am grateful for His mercy and the numerous blessings He has bestowed upon me.

I am grateful to my supervisor, Dr. Noor Farizah, for her consistent support, encouragement, and direction during my Ph.D. program. Dr. Noor has been an indispensable mentor and a continual source of motivation for me. Additionally, I express my gratitude to Dr. Mohammad Alauthman for his indispensable administrative assistance and contribution to data compilation and analysis.

I would also like to thank my family and friends for their love, support, and encouragement throughout my academic journey. My parents, my wife, Dr. Hind Alhamadeen, and my children's unwavering belief in me have motivated and strengthened me. I am grateful for their constant support and encouragement.

Finally, I want to thank all the participants who generously gave their time and shared their experiences for this research. With their participation, this research was possible.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

ACC	Accuracy
ACF	Autocorrelation Function
AdaBoost	Adaptive Boosting
ANTE	ANTicipating Botnets
AUC	Area Under the Curve
AutoML	Autonomous Machine Learning
C&C	Command and Control Server
CART	Classification and Regression Tree
CCC	Cyber Clean Center
CNNs	Convolutional Neural Networks
CRI	Compositional Rule of Inference
DALCNN	Live Capture Neural Network
DDoS	Distributed Denial of Service
DE	Differential Evolution
DNN	Deep Neural Network
DNS	Domain Name Service
DoS	Denial of service
DR	Detection Rate
DT	Decision Trees
EDIMA	Early Detection of IoT Malware Network Activity
FastGRNN	Fast-Gated Recurrent Neural Network
FIS	Fuzzy Inference Systems
FN	False Negative
FNN	Fast Neural Network
FP	False Positive

FPR	False Positive Rate
GA	Genetic Algorithms
GAN	Generative Adversarial Network
GBM	Gradient Boosting Machine
GLM	Generalized Linear Model
GRU	Gated Recurrent Unit
HRCA	Detecting High-Risk Coronary Artery Disease
HTTP	Hypertext Transfer Protocol
IDS	Intrusion Detection System
IF	Isolation Forest
IG	Information Gain
IIoT	Industrial Internet of Things
IoT	Internet of Things
IRC	Internet Relay Chat
ISP	Internet Service Provider
KNN	K-Nearest Neighbors
LiMNet	Lightweight Memory Network
LMT	Logistic Model Tree
LOF	local outlier Factor
LR	Linear Regression
LSTM	Long Short-Term Memory
LSTM RNN	Long Short-Term Memory Recurrent Neural Networks
MCC	Matthew's Correlation Coefficient
MCDA	Multiple-Criteria Decision Analysis
MLP	Multilayer Perceptron Neural Network
MNB	Multinomial Naive Bayes classifier
MOM	Mean of Maximum

OCC	One-Class Classifier
P2P	Peer-to-peer
PCA	Principal Component Analysis
RBMs	Restricted Boltzmann Machines
RF	Random Forest
RFE	Recursive Feature Elimination
RFIDs	Radio-Frequency Identification
RL	Reinforcement Learning
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
RTSP	Real-Time Streaming Protocol
SAW	Simple Additive Weighting
SDN	Software-Defined Networking
SVM	Support Vector Machine
TCP	Transmission Control Protocol
TN	True negative
TP	True positive
UDP	User Datagram Protocol

PENDEKATAN PELBAGAI PERINGKAT HIBRID BAGI PENGESANAN BOTNET IOT DAN PENILAIAN RISIKO TERTINGKAT

ABSTRAK

Percambahan peranti Internet Benda (IPB) telah membawa kepada cabaran keselamatan siber baharu. Isu penting ialah peningkatan kejadian Botnet IPB, yang merujuk kepada rangkaian peranti IPB yang terjejas seperti penghala, kamera IP dan peralatan pintar. Entiti yang terjejas ini digunakan secara strategik untuk menjalankan pelbagai ancaman siber, termasuk tetapi tidak terhad kepada serangan Penafian Perkhidmatan (DDoS) Teragih, penyingkiran data dan peninjauan rangkaian. Mengenal pasti Botnet IPB merupakan isu yang unik disebabkan oleh sumber yang terhad bagi peranti yang terlibat. Penyelidikan ini mencadangkan pendekatan berbilang peringkat yang menggabungkan susunan model ensemble dan peningkatan melalui sistem undian lembut untuk pengesanan IPB Botnet dengan memeriksa trafik rangkaian merentasi fasa pengimbasan, penyebaran dan serangan bagi kitaran hayat Botnet IPB. Penyelidikan ini menyumbang secara signifikan dengan mengenal pasti ciri penting yang dapat mengenal pasti setiap fasa dengan berkesan. Ciri-ciri ini termasuk penggunaan paket kawalan TCP yang ditapis, yang bukan sahaja mengurangkan trafik rangkaian tetapi juga meningkatkan prestasi pengesanan. Model ensembel berprestasi lebih baik daripada model individu, mencapai ketepatan pengesanan purata 91.7% bagi setiap fasa. Metodologi penilaian risiko berasaskan logik kabur yang dioptimumkan oleh kecerdasan kawanan zarah membolehkan penilaian fleksibel tahap keterukan serangan. Ia menyediakan asas kepada pasukan keselamatan untuk memperuntukkan sumber dengan cekap, membolehkan pertahanan keselamatan siber yang proaktif dan dinamik terhadap ancaman IPB Botnet. Kami juga

mencipta set data IPB yang realistik dan mewakili yang menyerupai kitaran hayat IPB Botnet dan termasuk serangan terbaharu dan canggih pada ekosistem IPB.

A HYBRID MULTI-TIER APPROACH FOR IOT BOTNET DETECTION AND ENHANCED RISK ASSESSMENT

ABSTRACT

The proliferation of Internet of Things (IoT) devices has led to new cybersecurity challenges. A significant issue is the increasing occurrence of IoT Botnets, which refers to networks of compromised IoT devices like routers, IP cameras, and smart appliances. These compromised entities are strategically utilized to carry out various cyber threats, including but not limited to Distributed Denial of Service (DDoS) attacks, data exfiltration, and network reconnaissance. Identifying IoT Botnets has unique issues due to the constrained resources of the devices involved. This research contributes significantly by identifying the active phase of the IoT Botnet attack life cycle and enabling flexible evaluation of attack severity levels through an ensemble model stacking and boosting via a soft voting system integrated with a fuzzy logic-based risk assessment methodology optimized by particle swarm optimization. This provides a basis for security teams to allocate resources efficiently, enabling a proactive and dynamic cybersecurity defense against IoT Botnet threats. A realistic and representative IoT dataset was also generated, simulating the IoT botnet lifecycle and incorporating the most recent attacks on IoT ecosystems. The proposed approach significantly advances IoT security by enabling precise detection of botnet activities and proactive threat mitigation. The integration of ensemble learning, fuzzy logic, and PSO offers a dynamic solution that adapts to evolving cyber threats, ensuring targeted, efficient responses and safeguarding network integrity.

CHAPTER 1

INTRODUCTION

1.1 Background

An attack is a malicious act that targets a system or network's vulnerabilities. Denial of Service (DoS) and Distributed Denial of Service (DDoS) attacks are sophisticated forms of cyber-attacks that can be carried out in several ways, including cracking passwords or intercepting network traffic. Online businesses and organizations face a growing risk from rising DDoS attacks, which can cause substantial monetary and operational losses.

A DDoS attack aims to impede legitimate users from accessing the targeted website or service by inundating it with traffic from multiple origins. Botnets, which are networks of compromised computers under the control of a single hacker, are commonly utilized to carry out such attacks. A Botnet is a group of computers or other devices that have been hacked and are used to perform malicious activities like spam, phishing, and widespread DoS attacks. In recent years, DDoS attacks that use Botnets have become more common. IoT devices are becoming more popular and easier to hack. Botnets can launch DDoS attacks of all kinds and types, from small attacks targeting single websites to large attacks that can take down whole networks. When Botnets are used, it is also hard to determine where the attack comes from because the information comes from many different places. A single compromised device can produce enough DDoS packets to launch attacks on such a massive scale that they could easily disrupt any IoT smart environment, as demonstrated in Figure 1.1, and destroy a functional IoT meta-system. BrickerBot Attacks (Cimpanu, 2017) and IoT-based Ransomware (Zahra & Shah, 2017) are examples of such attacks.

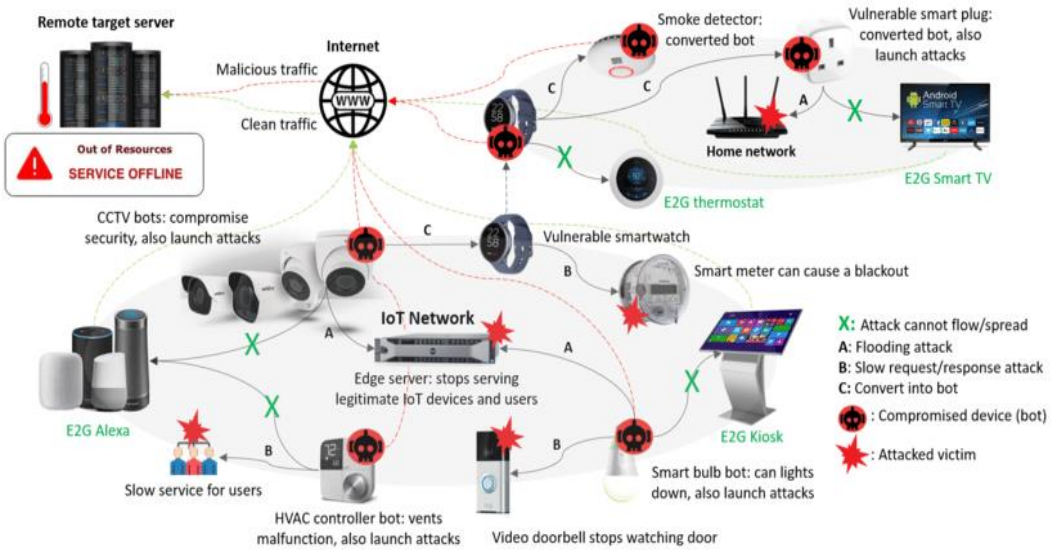

Figure 1.1 Attack Flow In A Typical Smart Environment (Sudharsan et al., 2021).

Using Botnets has made DDoS attacks harder to identify and mitigate, making them a danger to internet businesses and organizations. Organizations must adopt preemptive measures, such as integrating DDoS protection mechanisms, to safeguard themselves against such attacks. Moreover, it is imperative to enhance consciousness regarding the hazards linked with IoT devices and motivate users to implement requisite measures to safeguard their devices against potential breaches. Through the implementation of these measures, it is possible to make progress toward establishing a digital milieu that is both safer and more secure.

The IoT is a rapidly growing consumer and industrial technology. This technology can drastically impact the user experience by changing his daily lives, from how people drink coffee to how smart things interact with industry sectors. IoT could provide value in a variety of different ways. It can give businesses more detailed information about their organization, allowing them to enhance efficiency and cut costs. The term " Internet of Things " or "IoT" is becoming a popular buzzword among

all technology-related parties, including businesses and their customers. Technology changes people's lives in a big way and has a significant effect on the workplace. In which sensitive information is shared over the internet.

However, with the increasing number of IoT devices, the risk of cyber attacks targeting these devices has also risen. As a result, the attacker's curiosity can be turned into monetary gain. To achieve their objectives, attackers utilize numerous malware, such as viruses, worms, Trojans, ransomware, and adware. The Botnet is one of the most dangerous ways to perform malicious activities on the internet using various malware (Sicari et al., 2015). Over the years, firewalls and access controls have been implemented to ensure IoT devices are safe enough to overcome this problem. These methods focus on ensuring the data is private and authentic, controlling access within the network of IoT devices, and developing new security and privacy policies to help people trust each other.

Perpetrating DDoS attacks on IoT devices entails disseminating malicious software to many of these devices, thereby creating a network of bots. This network is subsequently utilized to attack a designated network or website. Frequently, the devices that have been infected need to be more adequately secured and left open to the internet, rendering them susceptible to attacks. These devices may become part of the Botnet without the owners' knowledge if they are infected with malware, which can be distributed via phishing emails or unprotected networks. Due to their limited processing power and memory, many outdated IoT devices remain exposed online. It happened because the security of the devices was not given enough thought to be included when they were produced at that time. Traditional cryptography is computationally costly for IoT devices, making them particularly vulnerable (Schöffel

et al., 2022). Other than that, weak login credentials are common on the IoT, whether pre-configured by the manufacturer or entered by the end user.

Although IoT devices have low processing and memory, there are ways to protect them from DDoS attacks. While it is true that devices can crash due to malware, there are many ways to protect user devices and prevent this from happening. The Botnet is being addressed to avoid flooding the network with traffic, allowing legal traffic to flow smoothly to all devices on the same network. Detecting Botnets may require powerful intelligence-gathering and data analysis from different sources, particularly backbone networks (Alauthman, 2016; Zhao et al., 2021). Intrusion detection approaches based on machine learning and deep learning have produced good results recently. An intrusion detection system (IDS) is designed as a cybersecurity tool that monitors a network or system for malicious activity or policy violations. It generates alerts when it detects suspicious activity and can be configured to take automated response actions (Almutairi et al., 2020).

The impact of DDoS attacks on IoT devices can be catastrophic. However, some measures can be implemented to mitigate the consequences. Effective security measures and vigilance can help mitigate such attacks' financial and operational repercussions and safeguard IoT devices. Despite such attacks, efforts have been implemented to prevent their recurrence, instilling confidence in these sites' continued provision of services to users. DDoS attacks utilize IoT devices' limited resources, such as storage and network bandwidth, and cause these problems in the IoT application (Al-Hadhrami & Hussain, 2021).

The safety and reliability of the internet and online services are significantly threatened by DDoS attacks targeting IoT devices. It is imperative to proactively safeguard IoT devices to mitigate the likelihood of such attacks. Several measures include utilizing robust passwords, consistently updating software, and establishing network segregation. Manufacturers of IoT products must prioritize security in their design processes. Secure communication protocols and firmware updates must be incorporated. Implementing these measures makes it feasible to ensure the sustained growth and advancement of the IoT while concurrently preserving its security and robustness.

1.2 Research Motivation

Working on early detection of IoT Botnets out of a sense of social responsibility or a desire to contribute to the greater good. By helping to protect against these threats, the detection may positively impact society and improve the overall security of the internet, improve cybersecurity, and protect against cyber threats more generally. Researchers have recently developed different methods for the early detection of IoT Botnet attacks. However, these developed detection methods covered the whole phase of a Botnet's lifecycle, namely scanning, propagating, and attack phase, making identifying the attack for each phase challenging as it is time-consuming to scan phase by phase (Wazzan et al., 2022). The challenge is identifying the specific IoT Botnet attack phase and warning the administrator of an impending attack. To the best of our knowledge, previous research has primarily focused on specific types or structures of botnets, often applying them to the entire lifecycle phase using limited or specialized datasets (Ahmad et al., 2022).

Besides, the reviewed studies observed that the ever-growing presence of IoT Botnets presents a major threat to vital systems such as electrical grids (Al-Turjman & Abujubbeh, 2019) and transportation networks (Oseni et al., 2022), requiring an effective strategy to handle and prevent suspicious activity. This challenge has made network administrators increase network traffic monitoring to identify potentially malicious activity from IoT devices. Based on this problem, early detection, as proposed in this study, is needed to help the administrators monitor and identify the potential attack specifically without having to check each phase, which will be time consuming and too late to launch an early control strategy.

With the advent of the IoT and the proliferation of linked devices, malicious actors' potential to initiate cyberattacks has increased. Therefore, businesses must create plans to rapidly and correctly detect Botnets to gain the upper hand over attackers before they can inflict major damage. Cyberattacks can cause significant financial and reputational damage to individuals and organizations. Cybercriminals constantly evolve tactics to bypass security measures, making preventing attacks challenging. Therefore, early detection is crucial in minimizing the impact of a cyberattack. Organizations can secure their assets and reduce the harm done by criminal actors if they plan to detect Botnets proactively.

The rapid growth of IoT networks has resulted in a large increase in linked devices and cyber risks, including IoT Botnet attacks. Insufficient security measures in these devices have rendered them vulnerable, with IoT devices accounting for 30% of all botnet attacks by 2023. The financial impact is significant, with IoT-related cybercrime costing $6 trillion worldwide in direct losses and brand harm. Over half of the attacks targeted critical infrastructure and enterprise networks. Experts anticipate

that by 2025, with 75 billion IoT devices, bot attacks will increase by 50% unless security is improved (Nasir et al., 2023). Furthermore, this emphasizes the critical need for increased security measures and additional research.

Recent cybersecurity reports from Cloudflare highlight the significant threat posed by IoT botnets, responsible for approximately 30% to 50% of all DDoS attacks. These reports indicate an alarming trend in exploiting poorly secured IoT devices, with notable incidents attributed to the Mirai botnet. Cloudflare emphasizes the evolving nature of these threats and urges organizations to adopt stronger security measures, such as enhanced authentication protocols and regular software updates, to mitigate the risks associated with IoT botnets. This underscores the urgent need for comprehensive strategies to protect vulnerable devices in an increasingly interconnected digital landscape [Aydn2021OCIDSAO] [McNulty2022IoTBC].

1.3 Research Problem

Early detection of IoT botnets remains a critical challenge in cybersecurity. Traditional detection methods often fail to identify botnet activities during the scanning and propagation phases, leaving networks vulnerable to attacks (Meng et al., 2017; Shtern et al., 2014).

Traditional detection systems that rely on single-agent methods or individual machine-learning models need help to deal with the complexity and diversity of contemporary data and threats (Catillo et al., 2023). The multi-tier approach uses multiple primary agents responsible for different parts of Botnet behavior to evaluate

and predict threats together. Early studies could have failed to adequately address the significance of detecting Botnets during their scanning and propagation phases (Alissa et al., 2022). These phases provide an opportunity to identify and isolate bots before they begin attacks such as DDoS attacks. Detecting scanning activities that occur early in an IoT Botnet attack is critical for identifying prospective Botnets and blocking future propagation. While many studies utilize application layer protocols such as DNS and HTTP for detection, these approaches often struggle to accurately identify IoT bots due to the unique characteristics of IoT devices and the evolving nature of Botnet attacks.

The lifecycle of IoT Botnet attacks is a complex process that involves multiple phases, from scanning, propagation, and attack phases. Detection of these attacks at early phases is crucial for minimizing their impact and preventing damage to connected devices and networks. However, traditional security solutions that rely solely on network layer protocols may not be effective in identifying these attacks, given the evolving tactics used by attackers and the diversity of IoT devices and protocols. Therefore, there is a need for a more comprehensive approach that leverages the application layer protocol to detect IoT Botnet attacks and identify the attack phase (Black & Kim, 2022). It will help to mitigate the risk of data breaches, service disruptions, and other negative consequences of these attacks.

Multiple studies have examined IDS performance in an IoT setting, with one report suggesting the use of ensemble-based deep learning models. These models frequently utilize sophisticated neural network structures, such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM), and Gated Recurrent Units (GRUs) (Lazzarini et al., 2023;

Mohamed et al., 2023; Odeh & Taleb, 2023). However, some limitations have been identified, such as high resource consumption, scalability issues, and performance challenges. Nevertheless, ensemble machine learning methods developed for IoT intrusion detection lack specialized agents dedicated to specific phase of Botnet operations (RQ2). This inadequacy presents a difficulty in precisely differentiating scanning, propagation, and attack activities inside the network.

Examining the subset of IoT features that play a fundamental role in identifying the attack phase Botnet attacks continues to be a research subject, which will reduce data dimensionality and facilitate learning and classification (Hikal & Elgayar, 2020). IoT devices have increased significantly in recent years, increasing IoT Botnet attacks. Identifying the attack phase can substantially aid in mitigating and reducing the impact of an attack. Thus, a comprehensive study is required to investigate the subset of IoT Botnet attack detection and identification features that are most effective. Reducing the number of features can result in several advantages, including a simpler and more manageable model, increased efficiency, improved interpretability, enhanced generalizability, decreased risk of overfitting, and decreased memory consumption. Selecting the right features can improve the accuracy of a model and reduce the time and resources needed to train it (Kelleher, 2019). A smaller set of features can make the model faster and more efficient and help find the most important factors for finding Botnets (RQ1).

A robust and innovative IoT Botnet detection framework that leverages a multi-tier approach will be developed to address these limitations. Utilizing a multi-tiered approach and ensemble learning strategies in IoT Botnet detection presents a promising solution. Hence, there is a pressing need to advance detection techniques

that are more efficient in targeting IoT Botnets' scanning and propagation phases. Traditional Botnet detection methods may have trouble distinguishing IoT device activity from Botnet activity (Nasir et al., 2023). The multi-tier approach may reduce false detections and improve accuracy by considering the stacking and boosting model outputs. The multi-tier aims to reduce false positives and improve Botnet detection accuracy. Integrating a voting mechanism for combining the results of the models mentioned above allows the proposed approach to leverage their respective advantages and offset any limitations, thereby enhancing the precision and efficacy of IoT Botnet identification.

IoT Botnet attacks are becoming more common as IoT devices become more commonplace, so it is important to have a systematic way to evaluate the potential for these attacks. Developing a risk assessment method for estimating the risk level of IoT Botnet attacks based on the attack phase's identification is one of the solutions to reduce IoT Botnet attacks. The current methods of only checking to see if an attack is happening still need to be improved (Kandasamy et al., 2020; Meidan et al., 2018). As a result, there is a requirement for a risk assessment technique that can precisely identify the attack phase and calculate the related risk level, hence facilitating more efficient mitigation measures as proposed in this study.

While current detection solutions primarily focus on identifying IoT Botnet attacks, there is still a need for responsive capabilities once threats are identified. Risk analysis is essential for properly assessing and responding to attacks to limit damage. Nevertheless, most existing systems tend to separate the detection and analysis components instead of integrating them into unified platforms. The absence of risk metrics delays the progress of proactive policy initiatives to strengthen defenses.

Implemented a Hybrid multi-tier approach that blends ensemble learning techniques with fuzzy logic to enhance the identification and evaluation processes, improving reaction agility and data-driven preparation. Integrating detection and analytics in this developing model showcases considerable promise for constructing a robust IoT ecosystem.

Existing IoT datasets, while useful, often have limitations in representing the diverse and evolving nature of IoT devices and attacks. Many datasets focus on specific devices or scenarios, making it challenging to develop comprehensive security solutions (RQ3). There is a need for more representative and up-to-date datasets that capture the full spectrum of IoT botnet activities across various devices and attack phases (Almaraz-Rivera et al., 2022; Lee & Ahmed, 2021). Using new datasets has multiple potential benefits for addressing IoT security and privacy issues. The key advantage is that it can aid in creating and precisely validating an IoT security solution (errag et al., 2022).

These challenges can be addressed by proposing a comprehensive and diverse IoT dataset that includes data from various IoT devices, sensors, and platforms, covers multiple application domains, and offers many features. The availability of a dataset of this nature would facilitate the development of more efficient security mechanisms and detection algorithms by researchers, improving the overall security and privacy of IoT devices and systems in diverse application domains. These new datasets will enable researchers to evaluate and improve the efficacy of intrusion detection systems, AI-based security solutions, and other IoT security mechanisms in fending off the most recent attacks. These datasets will advance IoT security research and improve IoT system security by capturing threats' dynamic nature and reflecting the evolving threat

landscape (Gyamfi & Jurcut, 2022). Developing new datasets is imperative for facilitating device identification and intrusion detection research, allowing for evaluating and enhancing IoT security mechanisms (Dadkhah et al., 2022; Damasevicius et al., 2020). It also contributes to advancing IoT security research by capturing the dynamic nature of threats and reflecting the evolving threat landscape (Gopalan, 2021).

A comprehensive and diverse dataset for IoT is necessary. Additionally, there is a need for a dataset that can simulate the three phases of the IoT Botnet Lifecycle: scanning, propagation, and attack. The need for comprehensive public datasets containing current traffic and attack patterns poses a significant challenge for developing machine learning systems in IoT environments. The existing special-purpose public datasets are inadequate in number, and the included attacking behaviors are outdated or insufficient, primarily due to the rapidly evolving threat landscape in the IoT domain (Sáez-de-Cámara et al., 2023). Creating a dataset that simulates the IoT Botnet Lifecycle can enhance the development of security mechanisms and detection algorithms, improving security and privacy for IoT devices and systems across various application domains.

This research aims to address these gaps by developing a hybrid multi-tier approach for early detection and enhanced the risk assessment of IoT Botnets, supported by a comprehensive and diverse IoT dataset.

This work intends to answer the following research questions:

RQ1. What features of network traffic flow can effectively be chosen as a subset to identify the IoT Botnet attack and its phase?

RQ2. How can an approach be developed to improve IoT botnet attack detection accuracy and reliability and enable risk-level assessment to facilitate appropriate mitigation actions?

RQ3. How can a realistic and representative IoT dataset be developed to capture the evolving IoT Botnet threat?

1.4 Research Objectives

This research aims to propose an IoT Botnet detection model capable of identifying the specific phase of the attack scanning, propagation, and attack phase for early detection based on application layer protocol.

To achieve the main goal, the following objectives have been formulated:

i) To identify the subset of significant features representing the IoT Botnet attack phase for a more efficient model.

ii) To propose a hybrid multi-tier approach that combines ensemble learning methods with an enhanced version of fuzzy logic using PSO to improve the accuracy of IoT Botnet detection, identify the attack risk level, and facilitate mitigation measures.

iii) To generate an IoT traffic dataset to facilitate advanced research in IoT security.

1.5 Research Contribution

The primary contribution of this research is a model to detect and identify the IoT Botnet attack phase. Following is a summary of this research contributions:

Theoretical Contributions:

i) Identifying a specific subset of significant and critical features for each phase of the IoT Botnet lifecycle to evaluate and determine the attack phase. Defining a subset of features to simulate each IoT Botnet's lifecycle phase to help assess and determine the attack phase. Defining a set of features for each phase of the Botnet lifecycle, such as the scanning, propagation, and attack phases. These features can be used to analyze the behavior of IoT devices and identify abnormalities indicative of Botnet activity. For instance, features like network traffic, device-to-device communication patterns, and software behavior could be relevant to the scanning phase. Network connections, data transmission patterns, and device-to-server interactions may be relevant in the propagation phase. Increase the accuracy and efficiency of IoT Botnet identification by focusing on a subset of relevant features for each phase of the Botnet's lifecycle.

ii) Developing a hybrid multi-tier approach that combines ensemble machine learning with an enhanced version of fuzzy logic using PSO to create a detection and analyzing model of the risks associated with IoT Botnets. The approach utilizes these two techniques to improve the accuracy and consistency of attack identification by utilizing the

strengths of diverse models. Moreover, integration enhances intelligent threat analysis by simulating human-centric reasoning for assessing risks on a graded scale. These insights can help security teams quickly assess the attack's severity and respond accordingly.

iii) Generating a realistic and representative IoT dataset containing the most IoT environment attacks and representing the attack's lifecycle phases. Such a dataset can help comprehend the complexities of IoT security, as it provides insights into the entire spectrum of attack behaviors, from scanning and propagation to full-fledged attacks. As IoT security advances, new datasets documenting recent attacks are needed. The proposed approach involves generating and disseminating new and inclusive datasets, which can catalyze the progress of research in the security domain for the IoT. Using this dataset, researchers can augment the advancement and assessment of intrusion detection methodologies and enhance the security of IoT applications and environments.

Operational Contributions:

In terms of operational contribution, an early discovery of the IoT Botnet can lead to more effective responses to lessen the harm of the attacks. It is known that when it comes to IoT Botnets, the scanning and propagation phase might take months (Kumar & Lim, 2019a). Hence, this research is designed to propose an effective solution to detect and isolate the bots early before a DDoS attack is launched.

1.6 Research Scope

This research focuses on IoT Botnet detection using Transmission Control Protocol (TCP) traffic within IPv4 environments, as IPv4 remains the dominant internet addressing protocol, handling over 90% of traffic. Limiting the scope to IPv4 allows for a focused analysis of the protocol most frequently targeted by cyber threats. Additionally, existing Botnet detection data, benchmarks, and frameworks are primarily based on IPv4, ensuring alignment with established methodologies and enabling valid comparisons (Shah & Parvez, 2014). The continued reliance of IoT devices on IPv4 further justifies this focus, as Botnets frequently target these devices. Prioritizing IPv4 also addresses the challenges many organizations face in migrating to IPv6, allowing the research to improve security measures for real-world IPv4 networks (Ashraf et al., 2023).

Since User Datagram Protocol (UDP) is a connectionless protocol, it does not provide the extra information needed to determine if a packet is a control or payload packet (Alauthman, 2016). As a result, the UDP packets are not considered in this research Table 1.1. List the overview of the scope involved in this research.

Table 1.1 Research Scope

No	Item	Scope of Research
1	Environment	IPv4
2	Attack type	IoT Botnet based on DDoS
3	Target layer	Application layer
4	Performance Metrics	Accuracy, False Positive, Detection Rate, Precision, Mean Square Error, Area Under the ROC Curve

Acknowledging the ubiquity and diverse impact of various cyber attacks such as spamming, credential stuffing, crypto-jacking, information theft, and privacy

breaches prevalent in IoT environments, this research navigates the IoT environment with a concentrated lens on DDoS attacks, one of the most pervasive types in the digital world (Srivastava et al., 2023). These attacks, particularly destructive due to their direct influence on service availability, can potentially invoke considerable financial and reputational damages. The focus on DDoS attacks is a strategic decision grounded in their broad ramifications, the existing vulnerability of IoT devices, and the profound societal implications should they be successfully executed. Instead, this research aims to look deeply at DDoS strikes in the IoT context. This effort is expected to add to the field of cybersecurity study and help make IoT environments safer by strengthening security protocols (Sommese et al., 2022).

The multi-tier approach proposed in the present research has the potential to function as an independent intrusion detection system for IoT environments. Organizations can implement the system on their networks to monitor traffic and receive notifications regarding potential botnet activity detected during scanning, propagation, and attack stages. The capability to analyze and get an insight into threats to the IoT could be further improved by integrating the proposed approach with current Security Information and Event Management (SIEM) systems. The risk score system and fuzzy logic engine provide enhanced capabilities for evaluating and addressing attacks. Integrating these analytics components into SIEM solutions will enhance their ability to make informed decisions regarding incidents related to IoT. In addition, managed security service providers can utilize customized versions of the multi-tier method to provide detection and response services for customers with IoT networks. Managed security services can utilize this capability to identify IoT botnets throughout client infrastructure and offer professional investigation and mitigation.

The scope of the research also includes creating a risk assessment model to enable the flexible evaluation of attack severity levels of the IoT botnet. To determine the severity of threats posed by IoT devices and to arrange responses accordingly, it integrates the results of multiple classifier agents with fuzzy logic. The PSO is utilized to optimize the parameters of the fuzzy system, including membership functions and rule weights, to improve its performance and adaptability. Integrating PSO optimization with fuzzy logic systems can automate the development process. The outcome of the fuzzy logic model categorizes the risk of an attack into various levels, including low, medium, high, and critical, depending on the severity of the threat and the potential impact on the IoT network.

The research utilizes benchmark datasets specifically designed for IoT Botnet detection, including CICIoT2023 (Neto et al., 2023), Bot-IoT (Koroniotis et al., 2019), MedBIoT (Guerra-Manzanares et al., 2020), and the Kitsune Network Attack Dataset (Mirsky et al., 2018). These datasets provide a comprehensive view of IoT network traffic, encompassing both malicious and benign data. CICIoT2023 offers a diverse range of IoT attack types such as DDoS, DoS, and spoofing, making it ideal for training detection systems with its real-world IoT environment representation. Bot-IoT features extensive records of both normal and malicious traffic, including detailed attack categories, which facilitates thorough evaluation of sophisticated botnet behaviors. MedBIoT focuses on the early phases of botnet deployment, including propagation and command-and-control communications, addressing critical stages often overlooked by other datasets. Finally, the Kitsune Network Attack Dataset is designed for lightweight intrusion detection, making it suitable for resource-constrained IoT environments and emphasizing anomaly detection through autoencoders. Collectively, these datasets offer a solid foundation for developing and validating the proposed

detection methods, covering a broad spectrum of attack scenarios and network traffic conditions.

1.7 Research Steps

This research examines different security approaches for the early detection of IoT Botnets. The literature review is the first step. This step begins with an investigation of the types of IoT Botnet attacks in general, followed by a study of the lifecycle of IoT Botnets and a review of previous work and related Botnet detection strategies in their early phases.

The analysis is the second step. This step begins with a study of the problems with existing IoT Botnet detection techniques. Next, approaches used within IoT Botnet detection based on Botnet lifecycle phases are evaluated to identify their limitations. Consequently, this work's problem statement outlines the proposed method.

Design and modeling assemble the third step. The approach is designed in this phase, and data preprocessing and feature extraction techniques are discussed. Additionally, the best subset of features is identified to identify the attack phase, and techniques for improving the detection accuracy are explored.

The Model Accuracy and Performance Evaluation is the fourth step. In this step, the experimental environment and implementation for assessing the efficacy of the proposed method and the accuracy of the model's results are described. This phase initiates with the execution of the methodology and culminates in evaluating performance metrics, including accuracy and false positive detection rate.

1.8 Thesis Organization

Chapter 1 introduction illustrates the research's background, motivation, problem, scope, goals, objectives, and contributions. This chapter also reviews the need for Botnet detection techniques and early Botnet detection.

Chapter 2 investigates the background of the research and related research. The techniques used to detect Botnets are discussed here, along with the advantages and disadvantages of each. This chapter also identifies the IoT Botnet detection research topic and introduces a new multi-tier approach for IoT Botnet detection. This chapter closes with a comprehensive discussion of the voids in existing techniques.

Chapter 3 describes the phases of the proposed approach, including data preprocessing, feature extraction, feature selection, and classification. The chapter also explains the proposed procedure for detecting the IoT Botnet phase in IoT network traffic.

Chapter 4 describes the research experimental work, containing the two experiments conducted. The chapter illustrates each experiment, including the setup, data collection, and results. The chapter concludes by discussing the results and their inferences.

Chapter 5 describes the dataset's creation, containing the experimental setup, data collection, and results. It also concludes by reviewing the results and their implications.

Chapter 6 summarizes the research's key findings, contributions, and limitations. It also discusses the research's implications and provides recommendations for future work. The conclusion chapter presents a comprehensive overview of the research and its significance in IoT Botnet detection and risk assessment.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter provides a complete and detailed background of IoT and Botnet detection. It also provides an overview of most methods used to discover IoT Botnets. It helps researchers understand the proposed research in Botnet detection and how to figure out the attack phase from the traffic characteristics and features.

2.2 Background of Study

2.2.1 IoT Architecture

Multiple IoT architectures have been proposed in the literature, such as SOA-based, middleware-based, and even six- and three-layer-based ones (Wazzan et al., 2021). In IoT architecture for essential communication, the three layers are a perception layer, a network layer, and an application layer, as follows (Raghuvanshi & Singh, 2020; Zhao et al., 2021):

The perception layer is a physical device and communication layer composed of sensors and actuators that absorb, sense, and manipulate data before sending it to the network layer. Cameras, RFIDs, and baby monitors are physical things in this layer. The network and transport layer are a communication layer that uses devices like gateways, switches, and routers to transfer and route consolidated data from the perception layer to the application layer. The application layer is a messaging layer that houses the application that communicates with users. Smart manufacturing, e-

health, and smart cities all rely on these applications. Illustrated in Figure 2.1 IoT Architecture.

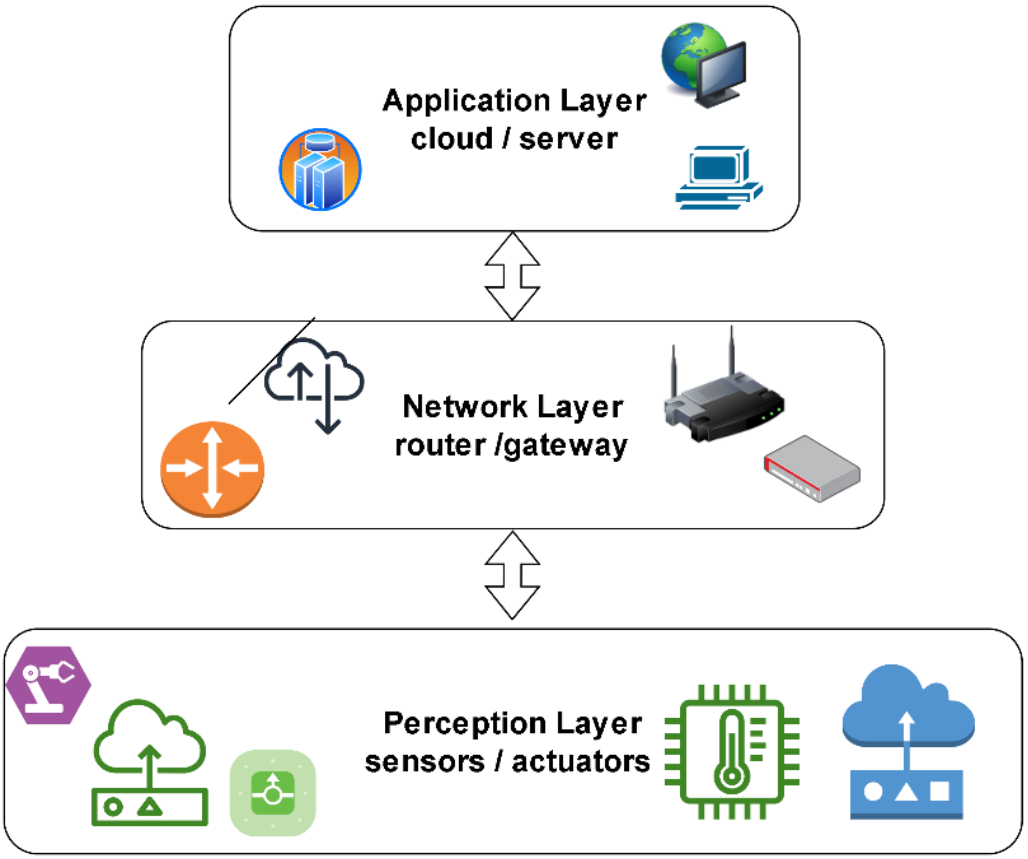

Figure 2.1 IoT Architecture (Peddolla, 2021)

2.2.2 Cybersecurity in IoT

IoT is one of the fastest-growing technologies, significantly impacting several fields (Gonzales et al., 2015). However, even though IoT has several benefits, one of the primary difficulties surrounding such devices is their enormous security weaknesses. According to a study headed by Hewlett-Packard, over 70% of the general IoT technologies have “some” vulnerability in their security, such as unencrypted data transmissions or excessively basic passwords (Kolias et al., 2015). IoT devices are one of the weakest links in the chain regarding securing a secure infrastructure since they

are susceptible to security flaws and are frequently deeply integrated into a network infrastructure.

Connectivity and diversity come hand in hand. IoT devices actively communicate data with each other. All these open links create several attack points. Participants' communication patterns, regulations, protocols, features, manufacturers, and security requirements vary (Zarpelão et al., 2017). Additionally, they are frequently geographically separated, meaning legislation from many countries may apply. As a result of their unique infrastructure, IoT networks are not well-suited to being protected by stringent cybersecurity procedures. Several groups, however, began formulating standards and recommendations that producers could utilize as a point of departure (Kolias et al., 2015).

Several studies (Anthi et al., 2019; Islam et al., 2020; Zarpelão et al., 2017) have found several reasons why IoT devices are vulnerable to hacking threats:

i) Heterogeneous devices increase IoT attack surface. Implementing security rules for ubiquitous IoT security is a tough challenge.
ii) IoT devices with limited processing power, memory, radio bandwidth, and battery life cannot perform security tasks sensitive to latency and require a lot of processing power.
iii) Most IoT devices operate without a human operator, making it easy for attackers to access them physically.
iv) It is challenging to apply patches or perform software updates on such devices.