A FEATURE SELECTION APPROACH BASED ON HYBRIDIZING FLOWER POLLINATION ALGORITHM WITH PARTICLE SWARM OPTIMIZATION FOR ENHANCING THE PERFORMANCE OF IPv6 INTRUSION DETECTION SYSTEM

ADNAN HASAN BDAIR AL GHURAIBAWI

UNIVERSITI SAINS MALAYSIA

2023

A FEATURE SELECTION APPROACH BASED ON HYBRIDIZING FLOWER POLLINATION ALGORITHM WITH PARTICLE SWARM OPTIMIZATION FOR ENHANCING THE PERFORMANCE OF IPv6 INTRUSION DETECTION SYSTEM

by

ADNAN HASAN BDAIR AL GHURAIBAWI

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

December 2023

ACKNOWLEDGEMENT

"I thank all who in one way or another, contributed to the completion of this research. First, I thanks "Allah" for the protection and the ability to do this research. I would like to give sincere thanks to the lecturers and collegemates at the National Advanced IPv6 Centre, the librarians, and other workers of the Centre for their kind help during my PhD journey, without you all, I will not be able to achieve this research".

"My special and hearty thanks to my research supervisor, Professor Dr Rosni Abdullah, and co-supervisor, Associate. Professor Dr Selvakumar Manickam. Without their assistance and dedicated involvement in every step throughout the process, this research would have never been accomplished. I would like to thank you very much for your support and understanding over these past four years".

"Getting through my research required more than academic support, and I have many, many people to thank for and, at times, having to tolerate me over the past three years. I cannot begin to express my gratitude and appreciation for their friendship. For many memorable evenings out and in, I must thank; Dr. Zaid Abdi Alkareem Alyasseri, Dr. Zakaria Alqattan, Dr. Amen Khallel Ibrahim and Dr. Shams alearifin have been unwavering in their personal and professional support during the time I spent at the USM university".

"Most importantly, none of this could have happened without my family. My life coach "My Father" Mr Hasan Bdair AL-Ghuraibawi, I really do not have any words to explain my thanks for your assistance to make me where I am now, without you, Dad, I am literally nothing. My lightness in this life is "My Mother," who encouraged me and prayed for me throughout the time of my research. And lastly,

thanks to my brothers Dr. Farhan AL-Ghuraibawi and Faleh AL-Ghuraibawi, Saleh AL-Ghuraibawi, Salman AL-Ghuraibawi, and Rehan AL-Ghuraibawi and Dr. Adel AIghuraibawi and my sisters and my love my wife Zahraa Muneam and my son Ahasan. I love you all".

"Finally, I would like to thank the unknown soldier who supports me without notice. May the Almighty God richly bless all of you. I dedicate this work to all of you".

Adnan Hasan Bdair AL-Ghuraibawi, Penang Malaysia, 2021

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

ACO	Ant Colony Optimization
ABC	Artificial Bee Colony
AC	Accuracy
ACDA	Auto-configuration Detecting Attacks
AI	Artificial Intelligence
AIDS	Anomaly-based Intrusion Detection System
Avg.	Average
BIAs	Bio-Inspired Algorithms
CPU	Central Processing Unit
DDoS	Distributed-Denial-of-Service
DoS	Denial-of-Service
DR	Detection Rate
FAR	False Alarm Rate
FPA	Flower Pollination Algorithm
FPR	False Positive Rate
GA	Genetic Algorithm
GNS3	Graphical Network Simulate 3
HIDS	Host-based Intrusion Detection System
ICMPv4	Internet Control Message Protocol version 6
ICMPv6	Internet Control Message Protocol version 4
IDS	Intrusion Detection system
IETF	Internet Engineering Task Force
IGR	Information Gain Radio

IPsec	IP Security
IPv4	Internet Protocol version 4
IPv6	Internet Protocol version 6
L2R	Local to Root
LA	Layered Approach
LAN	Local Area Network
LSSVM	Least Square Support Vector Machine
Max	Maximum
Min	Minimum
MITM	man-in-the-middle
MLD	Multicast Listener Discovery
MTU	maximum transmission unit
NA	Neighbor Advertisement
NDP	Neighbor Discovery Protocol
NFG	Network Flow Guard
NIDS	Network-based Intrusion Detection System
NS	Neighbor Solicitation
OS	Operating Systems
PCA	Principle Component Analysis
PSO	Particle Swarm Optimization
RA	Router Advertisement
RD	Redirect
RFC	Request for Comments
RS	Router Solicitation
SEND	Secure neighbor discovery protocol

SI	Swarm Intelligence
SIDS	Signature-based Intrusion Detection System
SVM	Support Vector Machine
TCP	Transmission Control Protocol
THC-IPv6	The Hacker's Choice IPv6
TNR	True Negative Rate
TPR	True Positive Rate
UDP	User Datagram Protocol
USM	Universiti Sains Malaysia

PENDEKATAN PEMILIHAN FITUR BERASASKAN HYBRID ALGORITMA PENDEBUNGAAN BUNGA DENGAN PENGOPTIMUMAN PARTIKEL SWARM UNTUK MENINGKATKAN PRESTASI SISTEM PENGESAN PENCEROBOHAN IPV6

ABSTRAK

Protokol Internet versi 6 (IPv6) adalah versi terbaru IP yang bertujuan untuk menampung beratus-ratus ribu alamat IP unik peranti dalam rangkaian. IPv6 mengemukakan ciri baru, seperti Neighbor Discovery Protocol (NDP) dan Skema Konfigurasi Auto Alamat. Untuk berfungsi dengan lancar, IPv6 memerlukan sejumlah protokol, termasuk ICMPv6. Ia dipikul dengan tanggungjawab utama. Serupa dengan IPv4, IPv6 rentan terhadah banyak serangan, termasuk jenis serangan yang lebih baru. Salah satu serangan berbahaya adalah serangan DDoS yang dilakukan melalui mesej ICMPv6. Serangan ini boleh memberi kesan keselamatan dan kewangan wang. Oleh itu, Sistem Pengesanan Pencerobohan (IDS) diperlukan untuk mempertahankan diri daripada serangan tersebut. IDS menggunakan ciri-ciri untuk mengesan serangan. Tetapi teknik pemilihan ciri, termasuk algoritma yang diilhami oleh bio, biasanya menghasilkan generasi subset ciri yang tidak sesuai, yang menghalang ketepatan pengesanan terhadap serangan DDoS yang dilakukan melalui mesej ICMPv6 semasa proses pembelajaran mesin. Sebilangan Algoritma Pencarian Pengoptimuman seperti itu terperangkap dalam minima carian tempatan, dan ini tidak pernah menganggap pendekatan pelbagai objektif, yang mengakibatkan pemilihan ciri yang tidak sesuai. Pilihan ciri yang tidak sesuai untuk IDS dapat diatasi dengan mengoptimumkan algoritma yang diilhamkan oleh bio dalam rangkaian IPv6. Dalam penyelidikan ini, Algoritma Pendebungaan Bunga (FPA) telah disesuaikan sebagai teknik pemilihan ciri

untuk mengenal pasti ciri yang paling relevan dari set data ICMPv6 untuk mengesan serangan DDoS ICMPv6 dengan tepat. Lebih-lebih lagi, hibridisasi algoritma FPA dengan algoritma PSO telah dicadangkan untuk meningkatkan ketepatan pengesanan. Selanjutnya, fungsi multi-objektif telah dicadangkan untuk bekerja dengan hibridisasi algoritma FPA dengan algoritma PSO untuk menilai dan mengurangkan bilangan ciri. Ciri-ciri yang dipilih digunakan untuk melatih set data menggunakan pengkelasan Mesen Vektor Sokongan (SVM). Pendekatan yang dicadangkan dinilai menggunakan set data ICMPv6 pada serangan yang berbeza. Hasil eksperimen menunjukkan bahawa pendekatan yang dicadangkan pertama mencapai ketepatan klasifikasi terbaik, iaitu, 97.96% dari segi jumlah ciri, dan ini mengurangkan jumlah ciri dari 19 menjadi 10 ciri. Di samping itu, hasil eksperimen menunjukkan bahawa pendekatan yang dicadangkan kedua mencapai ketepatan klasifikasi terbaik, iaitu, 97.99% dari segi jumlah ciri. Ini mengurangkan bilangan ciri dari 19 menjadi 8 ciri. Akhirnya, hasil eksperimen menunjukkan bahawa pendekatan ketiga yang dicadangkan mencapai ketepatan klasifikasi terbaik, iaitu, 97.01% dari segi bilangan ciri. Ini mengurangkan bilangan ciri dari 19 menjadi 4 ciri.

A FEATURE SELECTION APPROACH BASED ON HYBRIDIZING FLOWER POLLINATION ALGORITHM WITH PARTICLE SWARM OPTIMIZATION FOR ENHANCING THE PERFORMANCE OF IPV6 INTRUSION DETECTION SYSTEM

ABSTRACT

Internet Protocol version 6 (IPv6) is the most recent version of IP that aims to host hundreds of thousands of devices with unique IP addresses. In addition, IPv6 introduced new characteristics, including Neighbor Discovery Protocol (NDP) and Address Auto-configuration Scheme. For smooth functioning, IPv6 needed a number of protocols, including ICMPv6. It is vested with major responsibilities. Akin to IPv4, IPv6 is vulnerable to numerous attacks, including the newer type of attacks. One of the dangerous attacks is DDoS attacks carried out through ICMPv6 messages. These attacks could impose security and pecuniary implications. As a result, the Intrusion Detection System, also known as IDS, is required to provide protection against these kinds of attacks. IDS are constantly working on new features that will allow them to identify attacks. However, feature selection strategies, such as bio-inspired algorithms, typically provide an incorrect subset of features. During the process of machine learning, these characteristics hinder the detection accuracy of DDoS attacks utilizing ICMPv6 communications. Some of such Optimization Search Algorithms get trapped in local search minima, and these never considered the multi-objective approach, which resulted in an inappropriate selection of features. Such an inappropriate selection of features for IDS can be subdued by optimizing a bio-inspired algorithm in an IPv6 network. “The Flower Pollination Algorithm (FPA)” has been modified for use in this study as a technique for selecting features to determine which of the

“ICMPv6” dataset's properties are the most important in order to accurately detect “ICMPv6 Distributed Denial of Service (DDoS) attacks”. Moreover, hybridizing the “FPA algorithm with the PSO algorithm” has been proposed to enhance the detection accuracy. Furthermore, a multi-objective function has been proposed to work by hybridizing the “FPA algorithm with the PSO algorithm” to evaluate and reduce the number of features. The selected features are used to train the dataset using Support “Vector Machine (SVM) classifier”. The proposed approach is evaluated using the “ICMPv6 dataset on different attacks”. The experimental results show that the first proposed approach achieved the best classification accuracy, i.e., 97.96% in terms of the number of features, and it reduced the number of features from 19 to 10 features. In addition, the experimental findings demonstrate that the second proposed strategy achieved the best classification accuracy, i.e., 97.99% in terms of the number of characteristics. It reduced the number of features from 19 to 8 features. Finally, the experimental results showed that the third proposed approach achieved the best classification accuracy, i.e., 97.01% in terms of the number of features. It reduced the number of features from 19 to 4 features.

CHAPTER 1

INTRODUCTION

1.1 Overview

Nowadays, the Internet is ubiquitous and extensively used in nearly all facets of human life, such as businesses, industries, education sectors, entertainment, social media, health, and governance, spanning geographical regions, sectors, and fields (Fierro, Cardona Arbelaez, & Gavilanez, 2017). As a direct consequence of this, the Internet has developed into one of the most extensive man-made infrastructures and has emerged as an essential component of our day-to-day lives. The Internet helps users access information related to their interests anywhere and anytime. In addition, social networking brought people closer by collaborating and interacting remotely with others (Zhao, 2018; Sameera & Vishwakarma, 2019). Thus, the Internet has immensely supported us in many facets of our lives today more than ever. As a direct result of this, the number of people who use the Internet on a regular basis has increased substantially over the course of the previous five years, reaching an estimated 4.39 billion people around the world in 2016 (Reichelt, 2019).

Additionally, the number of cybersecurity issues has increased. It is becoming increasingly difficult as the number of Internet users and "Things" (i.e., Internet of Things) that must remain connected for an extended period of time increases (Elavarasi & Elango, 2017). Threat actors carry out cyber-attacks by leveraging vulnerabilities discovered in systems resulting in colossal loss of money. According to the report (BofA Merill Lynch Global Research), the estimated cost to the economy incurred by cybersecurity issues ranges from $539 billion to $1.46 trillion annually. Another report stated that Internet security issues could result in the extraction of value to the limit of

one-fifth created by the Internet in the potential worst-case scenario 2020 "Cybergeddon" (Symantec, 2015).

IPv4 is an abbreviation for "Internet Protocol version 4," which refers to the addressing protocol used by the Internet. The IPv4 address space has also been exhausted due to the rapid increase in the number and variety of devices that may connect to the internet. In addition to this, a whole new version of the protocol known as “Internet Protocol version 6” has recently been released (IPv6). It makes an effort to find a solution to the problem of IPv4 address exhaustion. The history of IPv6 is covered in the following part of this article. In addition, IPv6 has substitute the majority of its fundamental capabilities to the “Internet Control Message Protocol version 6 (ICMPv6)”, which is regarded as the primary protocol for any IPv6 network that is likely to function (Conta et al., 2006). “ICMPv6”, on the other hand, is affected by some implementation problems, including potential security risks, which need to be resolved. Therefore, a vast body of research indicates that ICMPv6 is susceptible to a variety of attacks that are already present in IPv4, including “Distributed Denial of Service “(DDoS) attacks and Denial of Service (DoS) attacks (DDoS)”. In addition to this, IPv6 introduces a new attack vector that relies on IPv6's additional properties, such as its ability to detect duplicate addresses, which were not present in IPv4 (Barker, 2013; Akamai, 2015).

The use of current technology, which can frequently have a significant impact on networks. This resulted in an increase in the amount of network intrusion attack that are carried out in the world we live in today against a variety of different sorts of networks. Integrating data analysis techniques, such as data mining and classification methods, with intrusion detection systems (IDSs) is a common practice that is undertaken to make IDSs more efficient. As a result, researchers have contributed, in

a variety of different ways, to the improvement of the use of IDS based on optimization algorithms work by extracting the most pertinent characteristics that can assist in accurately detecting network attacks (Sadiq *et al.*, 2018).

1.2 Background

Since the advent of the Internet, the Internet Protocol (IP) is the protocol that provides a unique identity to each Internet-facing node in the form of an IP address. However, the exponential growth of Internet-connected devices has exhausted all "Regional Internet Registries (RIRs) of allocatable IPv4 addresses" (Pickard *et al.*, 2017, 2018). Therefore, IPv6 has been engineered as the next generation of IP to resolve the issue. The Internet Society reported in 2018 that the usage of IPv6 had increased substantially (Internet Society, 2018). "Google revealed that IPv6" use had steadily increased from year to year, 6 per cent in 2015, 26 per cent in April 2019, and 33 per cent at the end of December 2020. Figure 1.1 depicts the percentage of users accessing Google services using IPv6 compared to IPv4 (Pickard *et al.*, 2019).

IPv6 was designed to improve the network's quality of service and security in comparison to IPv4 by reducing the complexity of routing and increasing its speed (Kristadi & Sucahyo, 2017). However, attempts at improving "IPv6 also introduce new vulnerabilities, such as with the Internet Control Message Protocol version 6 (ICMPv6)".

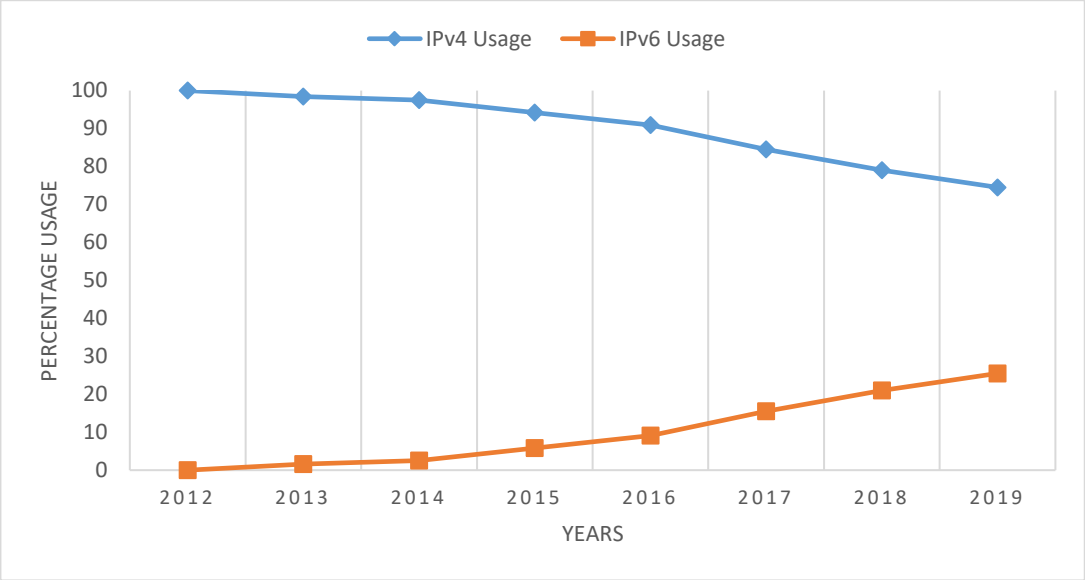

Figure 1.1 IPv6 Traffic Percentage Accessing Google Services

IPv6 adopts two standard mechanisms for configuring IPv6 addresses: stateful and stateless auto-configuration (SLAAC) (Ruiz *et al.*, 2017; Yousheng *et al.*, 2018). SLAAC relies on ICMPv6 messages through router solicitation and advertisement. Meanwhile, the stateful mechanism depends on Dynamic Host Configuration Protocol version 6 (DHCPv6) (Anbar *et al.*, 2018). The following section provides an overview of the many security concerns associated with the ICMPv6 Protocol.

1.2.1 Internet Control Message Protocol version 6 (ICMPv6)

Internet Control Message Protocol, as its name implies, is a protocol designed and used for control purposes, including testing, diagnosis, fault isolation, and reporting erroneous operations (Convery Sean, 2004). ICMPv6 is the implementation of ICMP for IPv6, which follows the same strategy as ICMPv4 (Saad *et al.*, 2013), except that it is an advanced version that plays a crucial role in IPv6 networks (Conta *et al.*, 2006).

IPv6 has many advantages over its predecessor, such as an autoconfiguration mechanism with the assistance of ICMPv6. The specifications of ICMPv6 are rife with significant modifications to the ICMPv4, such as the replacement of the "Address Resolution Protocol (ARP)" with the Neighbour "Discovery Protocol (NDP)" and changes in the administrative controls. NDP helps generate a node's unique IPv6 address without user intervention whenever a device connects to an IPv6 network (Narten *et al.*, 2007; Conta *et al.*, 2006). The NDP specifications separate the message types into two categories, i.e., error and information messages. The ICMPv6 messages fall under the category of information messages, where their messages comprise the type of message, checksum, and payload.

Table 1.1 shows the side-by-side comparison of ICMPv4 and ICMPv6 (Frankel *et al.*, 2010).

Table 1.1 Comparison of ICMPv4 and ICMPv6.

Distinctive	**ICMPv4**	**ICMPv6**
Next-Header (NH) value	(1)	(58)
control protocol	ARP	NDP
Increased Path Maximum Transmission Unit (PMTU)	576 bytes	1500 bytes
Multicast Listener Discovery (MLD)	IGMP has broadcast addresses	NDP does not have broadcast addresses

1.2.2 Security Issues in ICMPv6

Security flaws are apparent in the IPv6 network, and multiple vulnerabilities are related to ICMPv6. In addition, many researchers discovered various security threats on "IPv6, such as Distributed Denial of Services (DDoS)" and reconnaissance attacks (Durdağı & Buldu, 2010).

Another type of attack involves attackers sending many fake ICMPv6 messages to a targeted node on the network that forced the receiver to reply, increasing the node's burden, and resulting in unnecessary CPU utilization, culminating in performance degradation (Hogg S. Vyncke, 2008).

Another security issue involving the ICMPv6 protocol is it's affected by (DDoS) attacks, which could have disastrous impacts on IPv6 infrastructure, consume the victim machine's resources and bandwidth, and impact the network performance (Satrya *et al.,* 2015).

Another type of attack resulted in anomalous IPv6 traffic behaviour, which could severely impact the topology and properties of the network. Nevertheless, due to the increased need for filtering and inspecting in IPv6 networks, the abnormal behavior detection methods that are currently in use and built for IPv4 networks are not adequate for IPv6. Consequently, anomalous ICMPv6 traffic behaviour would remain undetected by such detection methods (Saad *et al.,* 2018).

(Elejla *et al.,* 2017) reported a vulnerability where multicast addresses might be used in a reconnaissance attack to identify or discover exploitable services and loopholes of the network system. ICMPv6 protocol is vulnerable to various attacks, including Distributed Denial of Service (DDoS) flooding attacks (Saad *et al.*, 2014).

"ICMPv6 protocol is vulnerable to various attacks, including DDoS flooding" attacks, becoming common Internet attack (Mowla *et al.*, 2014). Many organizations and Internet users suffered significant financial losses because of "DDoS attacks in the IPv6 network through ICMPv6 messages", representing a thorny problem and a severe issue of today's Internet. These attacks are prevalent due to the necessity of the ICMPv6 messages in an IPv6 network to function correctly *(Elejla et al.*, 2017).

DoS flooding attack is a common type of attack where attackers flood a network or a targeted node with enormous traffic. As a result, the targeted network or node becomes unavailable or remains out of service while forced to handle the massive network traffic (Saad *et al.,* 2014; Elejla *et al.*, 2017). Attackers could further modify the DoS attack to take the guise of a DDoS attack, whereby a router or a host was sent massive network traffic from multiple nodes to achieve the same evil objective. The basic principle of attack remains the same in "IPv6 networks as in IPv4 networks" (Alsadhan & Khan, 2013; Saad *et al.,* 2016). As stated in (Douligeris & Mitrokotsa, 2004; Safa *et al.,* 2008), a DDoS attack originates from many sources, contrary to a DoS attack. DDoS attackers almost always use spoofed source addresses to avoid tracking and detection (Douligeris & Mitrokotsa, 2004; Safa *et al.,* 2008). In addition, if attackers can establish unauthorized communication channels or slip maintenance messages into the link-local communication network undetected, they could invalidate legitimate addresses or disable interfaces, among other malicious acts (Saad *et al.*, 2016).

A reflector attack is an advanced form of a distributed denial of service attack, as seen in Figure 1.2. It comprises three main components, an attacker, an amplifying reflector, and a victim. The attacker spoofs its "IP address" with the victim node's IP address before sending the messages to the reflector machine. True to the role, the reflector nodes respond to the intended address, flooding the victim node with packets. The attacker node transmits "ICMPv6 echo" request packets by spoofing the victim's IP address to the multicast address of the network that will be amplifying the signal. It would appear that the victim node was the one that transmitted the echo request packets. All reflectors, excluding those configured not to respond to ICMPv6 multicast packets, would respond to each packet received. The reflectors unintentionally have

turned into an attacker by amplifying the effect. The reflectors were required to send packets at a reduced rate compared to the packet transmission rate of an ordinary DDoS attacker, who was supposed to directly flood the victim node (Beitollahi & Deconinck, 2012; Saad, 2016).

Another known attack that exploits the ICMPv6 message uses a Router Advertisement (RA) message. “Routers use RA messages” to announce their presence on a network segment. Hosts receive routing information and network prefix through such messages (Anbar *et al.*, 2018). RA messages, regrettably, were spoofed by attackers to swarm the victim node with unwanted traffic, thereby forcing the victim node to keep updating its neighbour cache table until exhausted. This process could be more potent using a DDoS attack, whereby fallacious RA messages appear to be sent by various routers through spoofing, making it even more challenging to detect (Elejla *et al.*, 2017).

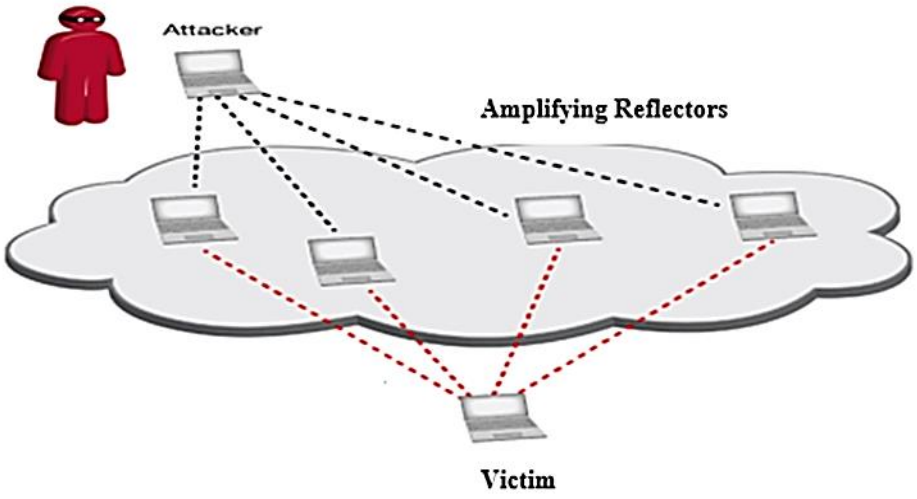

Figure 1.2 Architecture of DDoS Flooding Attacks Through Amplifying Reflectors

1.3 Intrusion Detection System (IDS)

Akin to new systems and protocols, IPv6 has also furnished requirements for thwarting attacks and imploring updates in protective mechanisms, including IDS (Caicedo *et al.,* 2009; Shah & Pervez, 2015). IDS is an intelligent tool designed to detect network intrusions and cyber-attacks by tracking and analysing customer devices or network traffic behaviour. Furthermore, it reports any intrusions or attack incidents detected to the relevant authority (Gholipour Goodarzi *et al.,* 2014).

Moreover, the IDS monitors ingress and egress traffic to and from different network links to detect unwanted events. On detecting an intrusion, it puts a record in the device log file for review by network security administrators so that the appropriate response or measure could be taken to protect the network. Such systems help detect intrusion and recover files if corrupted due to intrusion (Alomari & Othman, 2012).

It is worth noting that IDS built for the IPv4 network could also run in the IPv6 environment. However, it lacks the flexibility to cater to some IPv6-specific security problems, including “ICMPv6-based DDoS attacks”, due to the newly introduced specifications of “IPv6, such as NDP and auto-configuration mechanisms”. Moving ahead, several “IDSs have been developed to detect IPv6 attacks”; some of these were upgrades from existing IPv4-based IDSs (Elejla *et al.*, 2016).

Feature Selection Technique for IDS

According to (Bace, 2000) IDS is designed to monitor and analyze a large quantity of network traffic-related data. On the other hand, the massive amounts of data created by today's high-speed networks provide a challenge. Therefore, numerous types of intrusion detection systems run beyond the limitations of their processing capacities. Nevertheless, the current IDS is often characterized by a low detection rate

and a significant computational overhead owing to the need to process enormous packet lengths or data size, particularly in vast networks. Furthermore (Burggraeve, 2014; Hamdi, 2018), noted that network technology developments are making things more challenging. Therefore, it is of the highest necessity that the IDS process a limited amount of data to detect irregular behaviour in a real-time setting. (Saraswat Ayush, 2017) stated that for this reason, researchers frequently make use of a method known as "Feature Selection" to increase the effectiveness of “IDS” and the accuracy of detection. It is a strategy that allows for the selection or picking of a subset of mandatory features, which helps IDS accomplish its function. (Acharya & Singh, 2018) argued that, a feature subset that was chosen offers the possibility of improving testing and training, which would produce a reliable IDS model with high detection accuracy. Not only this, but this technique also reduces the complexity and computation time (Shen & Wang, 2012). Therefore, it is used to classify the more relevant features to achieve the objective of contributing to increasing the ICMPv6-based DDoS flooding attack detection rate. Selected features are assessed using an algorithm for learning calculation, wherein optimal feature subsets are derived. The precision of the classifier is very much crucial in choosing the right features. Even though the wrapper method is usually slower and requires extensive computation, it is still widely utilized due to its acceptability (Kumari & Swarnkar, 2011).

1.4 Research Motivation

The IPv6, with its many new features, also suffers from security vulnerabilities, which hampers its deployment. For example, DDoS, is a common attack in IPv6 networks, where attackers exploit the security loopholes in ICMPv6 and NDP to attack the victim’s machine (TechCenter, 2013). This alarming situation motivates

researchers to propose new methods and technologies to detect unknown and novel "DDoS attacks based on ICMPv6 messages". The motivations for this research are hereunder:

1. An IPv6 Cybersecurity Surveillance Survey report in 2014 looked at respondents' security concerns from Arbor Network Inc. The report showed that DDoS attacks are the main cyber threats overall for IPv6 networks. The growing phenomenon of DDoS attacks to difficult to detect at 52% of respondents involved compared to 72% the previous year (Alangar and Swaminathan, 2013, Arkko *et al.,* 2011, Ektefa *et al.,* 2010, Radhakrishnan *et al.,* 2007, Saad *et al.,* 2016).
2. IPv6 is currently protected by a great number of security methods; nevertheless, these mechanisms are not yet at the required level when compared to "IPv4 in terms of their capabilities and performance" (Gont, 2011, Gont, 2012a; SAAD, 2016; Elejla *et al.*, 2017, 2018).
3. The "Internet Control Message Protocol version 6, or ICMPv6", is the most essential protocol related to the IPv6 protocol, and it now has an increased function, particularly in the auto-configuration mechanism. In addition to this, "ICMPv6 messages" are split into two categories: those that are deemed NDP messages belong to the category of informative messages. The security precautions that were taken with ICMP version 4 are no longer adequate to deal with the possible threats that are presented with ICMP version 6. Therefore, the lack of security consideration in the design of the protocol led to several vulnerabilities, such as a potential attack vector that simply involves delivering a large number of unlawful "ICMPv6 messages" to a network device, which will result in the load on the node is increased. However, this circumstance

could lead to increased CPU consumption, which would result in a decline in performance. Another "vulnerability of ICMPv6" is that it can be used by an attacker to carry out attacks of (DoS and DDoS)(Shanmugaraja and Chandrasekar, 2012, Oliveira *et al.,* 2012; Elejla *et al.,* 2016; Saad, 2013, 2016).

Flower Pollination Algorithm (FPA) (Yang, 2012) is a population-based algorithm developed in 2012 with two operators, i.e., local and global pollination operators. However, the FPA has not been used as the search algorithm in any existing IPv6-based detection system to the best of our knowledge. Therefore, this research focuses on adapting FPA for detecting ICMPv6-based DDoS attacks and then enhancing its efficiency by hybridizing with other meta-heuristic algorithms and multi-objective functions.

1.5 Problem Statement

Recently, online services have witnessed rapid expansion in Internet services. But it is vulnerable to intrusions from malicious content. Therefore, researchers continuously seek measures and techniques to protect online services and communication networks from intrusions, such as DDoS flooding attacks through ICMPv6 messages (Saad, 2016). It is well-known that techniques for detecting common and existing attacks probably fail in detecting new attacks, such as "ICMPv6-based DDoS" flooding attacks. Malicious packets during DDoS attacks may exhaust the resources of the targeted CPU and degrade performance (Elejla *et al*., 2017; 2018).

A computer network transmits an enormous amount of data that must be devoid of malicious content or behaviour that degrade network performance or threaten users' privacy. One tool that analyses data traffic to detect malicious traffic behaviour is

Intrusion Detection System (IDS). IDSs commonly use Meta-heuristic algorithms to recognize cyber-attacks, including ICMPv6 DDoS flooding attacks. However, prevalent methods lag in using appropriate features selected from a dataset to detect flooding attacks (Elejla, 2018).

Increasing the detection accuracy of the IDS requires analysis features. Various filter-based algorithms, such as Principal Component Analysis "(PCA) and Information Gain Ratio (IGR)", are useable for analysis features. However, these algorithms choose a subset of features that could affect detection accuracy and the classifier's performance (Kuang, Xu, & Zhang, 2014).

Some researchers hybridized the wrapper and filter algorithms to achieve the best subset with reasonable accuracy. It is also vital to reduce the complexity and computation time. Moreover, some researchers combined filter algorithms with Meta-heuristic algorithms to select features. But the results did not reveal good performance (Ganapathy *et al.*, 2016).

Meta-heuristic algorithms are primarily used in optimization and pattern recognition scenarios. These algorithms offer diversity and flexibility to respond on the spur of the moment to unknown situations. Thus, they are suitable in the IPv6 network for countering unknown attacks (Zulkiflee *et al.*, 2015). At the same time, such meta-heuristic algorithms face difficulties in maintaining the equilibrium between global and local search exploration and thereby avoiding being trapped in the local search minima. Moreover, swarm-based or population-based algorithms tend to have the capability to search across several regions in the problem landscape (Abdi *et al.*, 2018).

(Jeyavim Sherin and Parkavi, 2022) stated that in recent years, meta-heuristic algorithms have emerged on the scene. These metaheuristics are designed to simulate the process of solving optimization issues. In addition, bio-inspired algorithms are becoming increasingly popular, which is bringing about a revolution in the field of computer science. carries out research into the method of feature selection utilized by meta-heuristic algorithms -driven intrusion detection systems. classifies different SI techniques according to the many parts of an intrusion detection process in which they are applicable and how they might improve those aspects. (Hosseini, Gharehchopogh and Masdari, 2022) argued that a large number of academics have investigated how the detection performance of datasets changes when the number of features in those datasets is reduced. One kind of data mining that is successfully included in detection and identification systems that are built for attacks is the selection of numerous attributes from a dataset. Accordingly, the amount of features that are chosen from the dataset and the relevancy of those features can both have a direct influence on the detection error rate. As a result, we used a brand new multi-objective detection technique that is both practical and efficient, and it is based on feature selection.

Researchers have developed a wide variety of meta-heuristic algorithms, and these algorithms have been implemented in a wide variety of domains with outstanding results. Despite this, not a lot of implementation work has been done in the IDS feature selection area. In addition, the majority of the algorithms use the detection accuracy rate when performing the evaluation, which helps to improve the IDS's overall performance. On the other hand, did not take into account the number of features that were being used, which led to an increase in complexity, utilization of system resources, and amount of time spent computing (Gholipour Goodarzi *et al.*, 2014). In addition to this, the multi-objective function is sometimes referred to as the weight

sum fitness function. The number of chosen characteristics will be narrowed down with the help of this fitness examination. While doing so, achieve a high level of accuracy in classification. The appropriateness of the solution is judged according to the degree to which it presents the specified quantity and type of selected features accurately. Each candidate solution includes a set of chosen characteristics. If two candidate solutions with differing numbers of selected characteristics get the same value for accurate detection. The candidate with less features has been selected by the algorithm, and the other candidate will be disregarded (Alamiedy *et al.*, 2019).

The main research questions identified for this research are in the following points:

1. Can a meta-heuristic algorithm be adopted to find an optimal subset of features with a high accuracy rate to detect the "ICMPv6 DDoS" attack messages?
2. Can the standard meta-heuristic algorithm be customised to find an optimal subset of features with a high accuracy rate to detect the "ICMPv6 DDoS attack?
3. Can a multi-objective function be explored to obtain the best feature to detect the ICMPv6 DDoS attack messages and improve accuracy?

1.6 Research Objectives

This research aims to improve the IDS performance by optimizing a bio-inspired algorithm in an IPv6 network encountering anomalous behaviour to identify ICMPv6 DDoS flooding attacks. The research objectives are as follows:

1. To adapt a meta-heuristic algorithms algorithm to identify an appropriate subset of features that increase classifier accuracy to detect “ICMPv6 DDoS flooding attacks”.
2. To hybridize the Flower Pollination Algorithm with another meta-heuristic algorithm to select optimal features subset for improving the classifier accuracy to detect “ICMPv6 DDoS flooding attacks”.
3. To propose a multi-objective function to improve the selection of features by the hybrid meta-heuristic optimization algorithm aims to minimize the number of features and improve detection accuracy.

1.7 Research Contribution

The primary contribution of this research is the extraction of an optimal subset of features through an optimization algorithm to detect DDoS flooding attacks by monitoring anomalous behaviour of ICMPv6 traffic in IPv6 networks to achieve a high detection rate. Furthermore, the optimal subset of features extracted will be used to train the machine learning classifier model through a meta-heuristic algorithm. The sub-contributions are summarized below:

1. “A Flower Pollination Algorithm” is used as a feature selection approach to detect DDoS attacks based on the ICMPv6 DDoS attack dataset.
2. Hybridization of the Flower Pollination Algorithm with Particle Swarm Optimization (PSO) improves feature selection for the ICMPv6 DDoS attack detection.
3. A multi-objective function approach selects the optimal features subset to improve and maximize detection accuracy and minimize the number of features.

1.8 Research Scope and Limitations

The experiments conducted in this thesis are limited to detecting DDoS attacks that use ICMPv6 messages in IPv6 networks. ICMPv6 works at the network layer; therefore, attacks on other layers are beyond the scope of this research. The proposed approach works on an anomaly-based IDS; consequently, it does not maintain any database for attack signatures. Optimum subsets obtained are packet-based datasets containing malicious traffic generated through The Hackers Choice's (THC) attack toolkit. The detection accuracy rate of targeted attacks is employed as the evaluative metric in this situation. The extent of this research as well as its limitations are detailed in Table 1.2.

Table 1.2 The Research Scope and Limitations

Item	Scope of Research
Environment	"Internet Protocol version 6 (IPv6) network"
Detection	Anomaly-based detection
Targeted Layer	Network layer
Attack type	RA and NS DDoS flooding attack
Protocol	"Internet Control Message Protocol version 6 (ICMPv6)"
Dataset	Normal and Abnormal ICMPv6 Traffic Packet-based Dataset
Attacking Tools	"The Hackers Choice's (THC)-IPV6 Attack Toolkit"
Evaluation	Detection accuracy rate and number of features

1.9 Research Steps

The solution for increasing the DDoS attack detection accuracy is by monitoring the anomalous ICMPv6 traffic behaviour in an IPv6 network using a meta-heuristic algorithm as a method for selecting features from a dataset in order to select the best possible subset of those features. Practical steps are taken to achieve the goals of this study. The first step is to review related literature and analysis therein. The

second step proposes a new mechanism to detect “ICMPv6 DDoS flooding attacks in an IPv6 network”. The third step is to design and implement the proposed mechanism. The fourth and final step is to test and evaluate the results and present the research findings. Figure 1.3 illustrates the methodological steps undertaken in this research. The **first phase** involves critical reviews of the existing studies. This phase aims to comprehend the problem and future scope of research to detect DDoS flooding attacks that use ICMPv6 messages in an IPv6 network with increased accuracy.

The solution to the problem ascertained in the first phase is presented in the **second phase**. The solution comprises specific stages to detect DDoS flooding attacks by monitoring anomalous ICMPv6 traffic behaviour in IPv6 networks with increased detection accuracy. The mechanism is proposed; wherein, features are optimally selected using an optimization algorithm. IDS is meant to monitor and analyze a vast number of data connected to network traffic. The huge amounts of data generated by today's high-speed networks, however, present a dilemma. Consequently, a variety of intrusion detection systems exceed their processing limits. A hybrid FPA algorithm shall be a foundation for detecting attacks with increased accuracy.

The **third phase** involves designing the proposed mechanism and its implementation to increase the detection accuracy of “DDoS flooding attacks based on ICMPv6 messages in an IPv6 network”. The crux shall be the efficacy in detecting intrusion, selecting features, and training the model with appropriate input.

The test and evaluation shall be carried out in the **fourth phase**. The proposed mechanism comprising a “hybrid FPA algorithm amalgamated” with a meta-heuristic algorithm will be tried and appraised for its efficacy in “detecting DDoS flooding attacks based on ICMPv6 messages in an IPv6 network” on an actual stream of data

created from a testbed. Finally, the proposed mechanism will be compared with the existing anomaly-based mechanisms for "detecting ICMPv6 DDoS flooding attacks".

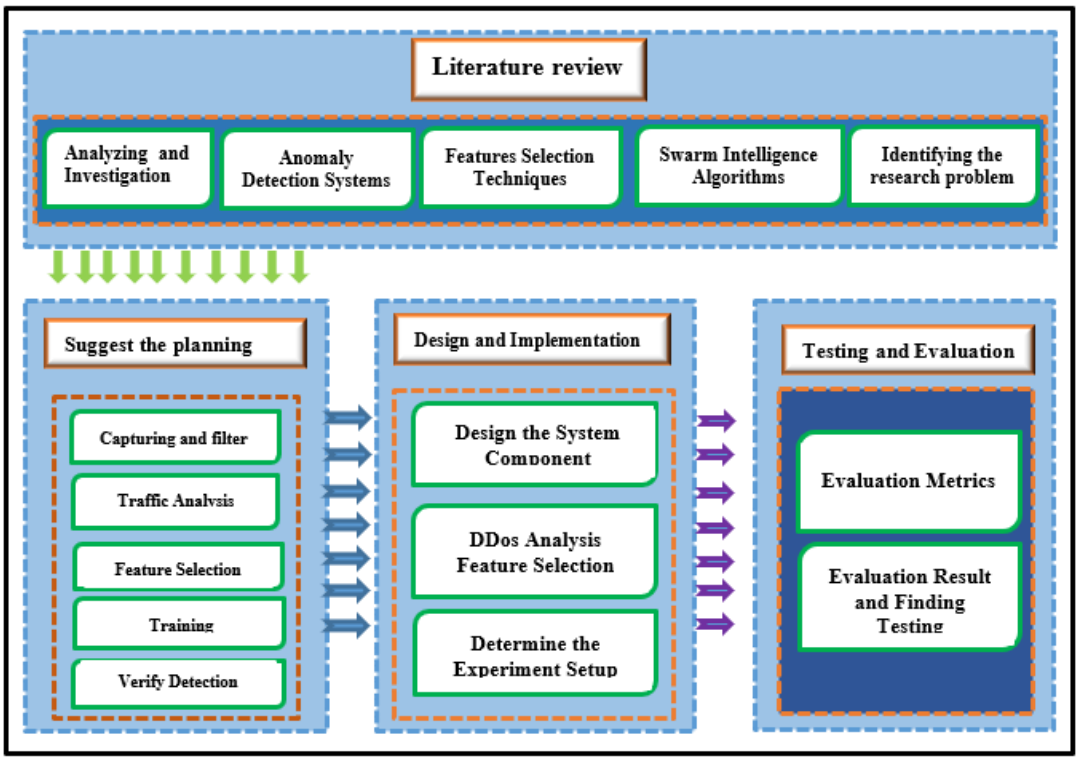

Figure 1.3 Research Steps

1.10 Thesis Organization

This thesis has six chapters.

Chapter 2 provides an overview of the security aspects, a description of features pertinent to the ICMPv6 flooding attack, a review of the existing literature and research under the axis of this study, and an overview of the detection mechanism to be utilized as a basis of the framework.

The proposed methodology is in **Chapter 3**; wherein, an elaboration of the proposed solution and constituent segments of the proposed algorithm are presented.

Chapter 4 demonstrates the design of the proposed mechanism, along with its implementation in three stages. (i) collection of data and data pre-processing; (ii) preparation of the "ICMPv6 message dataset"; and (iii) adapting "FPA" as a feature selection technique, which aims at "detecting ICMPv6 DDoS flooding attacks in an IPv6 network". "A hybrid FPA amalgamated" with an optimization algorithm is proposed that aims to improve detection accuracy. In addition to this, a multi-objective function is used to identify the ideal set of features that can contribute to high detection accuracy. This is done by comparing the importance of each feature to the importance of the others. This is done in order to maximize the efficiency with which the detecting system operates.

Chapter 5 provides evaluation and testing of the subset of data and deliberation on the proposed approach concerning the IDS. The results obtained from detecting "ICMPv6 DDoS flooding attacks in an IPv6 network using FPA" are discussed. Moreover, it also discusses the results of hybridizing the FPA with an optimization algorithm and multi-objective function to select an optimum subset of features for the detection of "ICMPv6 DDoS flooding attacks in an IPv6 network" with high detection accuracy. At a later step, the suggested mechanism is tested and compared with existing methodologies with regard to the measure of attack detection accuracy.

Chapter 6 presents the research contributions, the conclusion, and recommendations and future works.

CHAPTER 2

LITERATURE REVIEW

This chapter covers the related literature on anomaly-based intrusion detection techniques for detecting “DDoS attacks based on ICMPv6 messages in an IPv6” network. The chapter has eight sections. Section 2.1 presents the backdrop of the research. Section 2.2 explains the improvements in IPv6, the impacts on security due to such improvements, the importance of “ICMPv6, and security issues in IPv6”. Existing IPV6 Datasets are reviewed in Section 2.3. Types of intrusion detection systems for ICMPv6-DDoS attacks are reviewed in 2.4. The feature selection techniques are in Section. 2.5, followed by a discussion of the related work in Section 2.6. Section 2.7 visualizes the research gaps, and the last section summarizes this chapter.

2.1 Introduction

Since its inception, IPv6 is prone to vulnerabilities. To make it secure, network administrators need to place appropriate security policies by assessing possible risk factors according to network requirements (Rosli *et al.*, 2018).

Securing an IPv6 network is compulsory because IPv6 is inherently insecure. Ultimately, it is replaced by IPv4, therefore, to ensure that the IPv6 network connection between the node is secure, DDoS attacks attempt to disrupt the network services of the target device, causing the potential failure to send “IPv6” packets. In response, several researchers have sought to find a methodology to mitigate and detect “DDoS attacks”. Figure 2.1 illustrates the backdrop of this research.

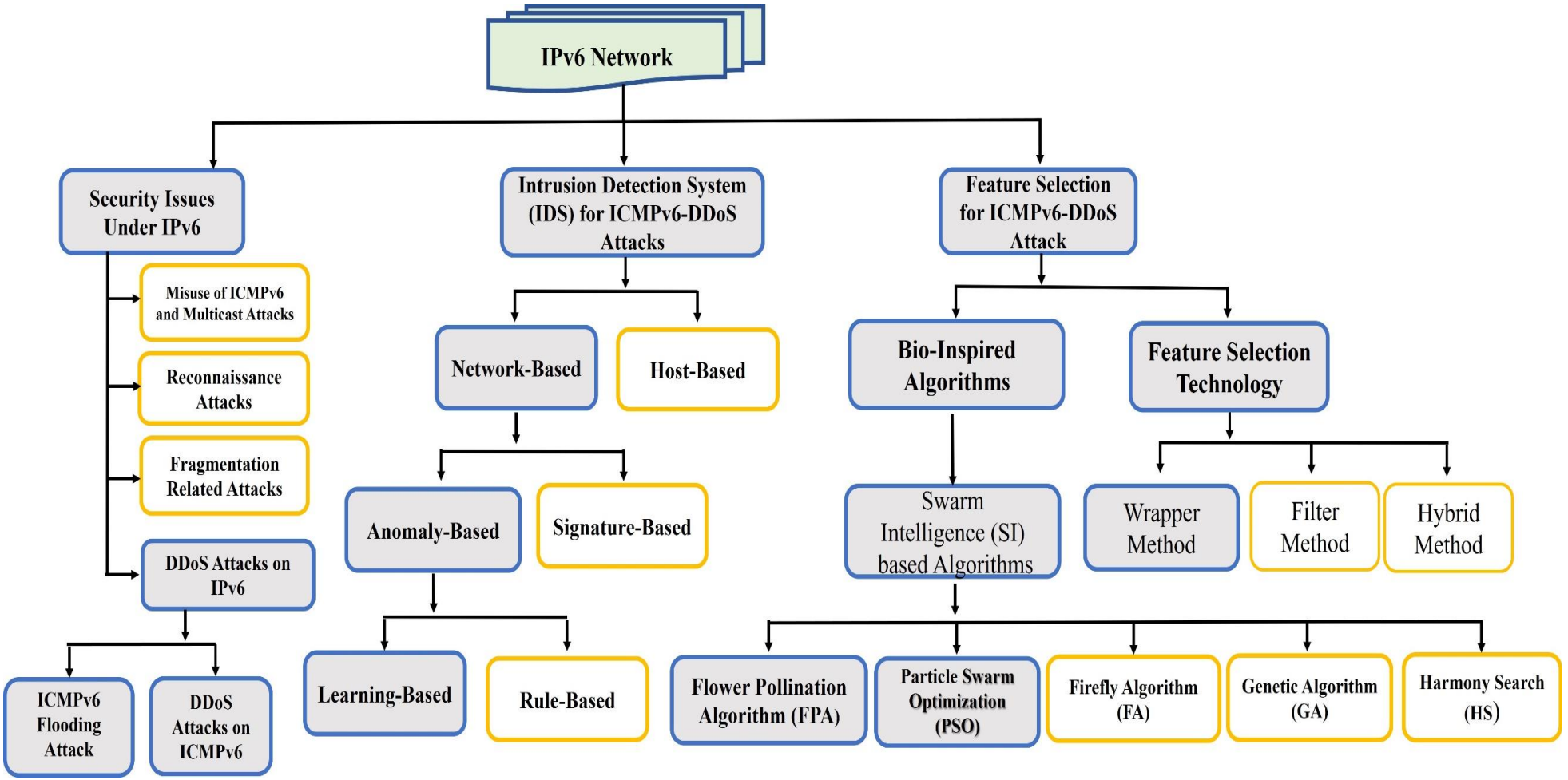

Figure 2.1 The Backdrop of Study

2.2 Internet Protocol version 6 (IPv6)

Internet users are increasing with every passing day. Hand-held devices have significantly revolutionized the Internet in this regard. There arose a need for an enormous amount of additional IP addresses, which was fatal for IPv4 (Deering, Fenner, & Haberman, 1999). Therefore, the engineers proposed IPv6 to resolve the address exhaustion problem in 1998 (Elejla *et al.*, 2017; Al-Ani *et al.*, 2019). IPv6 can support a practically unlimited number of devices (Al-Ani *et al.*, 2019). IPv6 was launched on June 06, 2012, by the Internet Society (ISOC). On March 09, 2020, the "proportion of clients using Google services over IPv6 exceeded 25.88% to 32.62% on 7 October 2021", and the percentage has been steadily increasing since then, as shown in Figure 2.2 (*IPv6 – Google*, 2021). The trend shows that IPv6 is likely to be users' choice in the not-so-distant future.

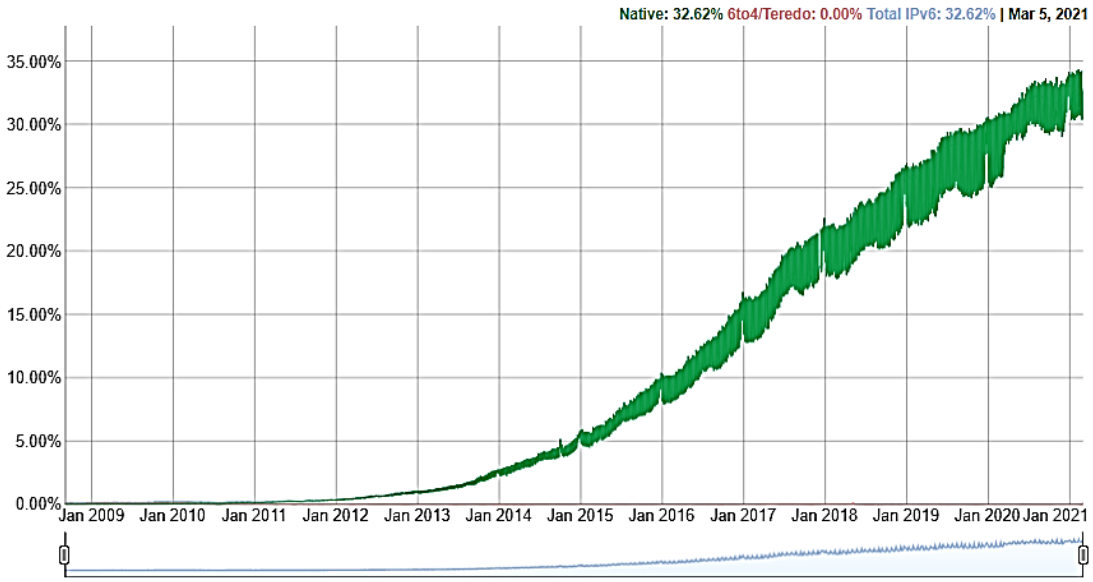

Figure 2.2 Google Service-Connected Through IPv6

IPv6, like other protocols, has a format that includes a fixed header. In IPv6, Extension Headers replace IPv4 'Option' type field. These Extension Headers are flexible as they are not part of the main header and include Authentication Header, Encapsulation Security Payload, and so on (Crainicu & Maior, 2005).

Nevertheless, IPv6 has inherited several IPv4 features (Supriyanto *et al.*, 2013; Elejla *et al.,* 2017). However, IPv6 security is more robust than IPv4, and the credit goes to IPsec. It unleashes auto-configuration, mobility, and extensibility. NDP is an auto-configuration protocol capable of generating IPv6 addresses for IPv6-capable devices once they connect to the network (Narten *et al.*, 2007). IPv6 nodes depend on the ICMPv6 protocol to detect a new device, neighbour, and router in the network. The ICMPv6 protocol also provides diagnostic and error reporting functions to IPv6 networks (Conta *et al.*, 2006).

2.2.1 Internet Control Message Protocol version 6 (ICMPv6)

IPv6 comprises various IP suites, including the ICMPv6 protocol (Postel, 1981). ICMPv6 is the backbone of IPv6. It copies the strategy of ICMPv4 in testing, controlling, and sending error messages back to the originating IP address *(Saad et al.,* 2013). Type value ranges between 0 and 255 in ICMPv4 and ICMPv6. However, ICMPv6 message types are different from ICMPv4 (Conta *et al.*, 2006), although both value and message types are unbonded. ICMPv6 specifies 0 to 127 as the values for error messages and 128 to 255 for informational messages. The value for the Next "Header in the IPv6 packet is 58 for ICMPv6 messages". Table 2.1 shows the comparison of ICMPv4 and ICMPv6 messages (Saad *et al.,* 2013).