

# **ESCUELA POLITÉCNICA NACIONAL**

**FACULTAD DE INGENIERÍA DE SISTEMAS**

**LEARNING MODELS TO DETECT PERSONALITY TRAITS OF  
CYBER ATTACKERS: A COMBINED APPROACH USING  
HONEYPOT AND SURVEYS**

**THESIS SUBMITTED AS PART OF THE  
REQUIREMENTS FOR THE AWARD OF THE DEGREE  
OF DOCTOR OF PHILOSOPHY IN INFORMATICS**

**CARLOS JOSÉ MARTÍNEZ SANTANDER**

carlos.martinez03@epn.edu.ec

**ADVISOR: HUGO OSWALDO MORENO AVILÉS**

h.moreno@epoch.edu.ec

**COADVISOR: MYRIAM BEATRIZ HERNÁNDEZ ÁLVAREZ**

myriam.hernandez@epn.edu.ec

**Quito, July, 2024**



ESCUELA  
POLITÉCNICA  
NACIONAL

## THESIS

For the award of the degree of

## DOCTOR OF PHILOSOPHY IN INFORMATICS

Resolution RPC-SO-43-No.501-2014 of the Consejo de Educación Superior

Presented by

**CARLOS JOSÉ  
MARTÍNEZ SANTANDER**

Thesis supervised by

**Hugo Oswaldo Moreno Avilés,**

**Professor of the Escuela Superior Politécnica of Chimborazo  
(Ecuador)**

and co-supervised & Tutor by

**Myriam Beatriz Hernández Álvarez,**

**Professor of the Escuela Politécnica Nacional(Ecuador)**

**LEARNING MODELS TO DETECT PER-  
SONALITY TRAITS OF CYBER AT-  
TACKERS: A COMBINED APPROACH  
USING HONEYPOT AND SURVEYS**

Oral examination by the following committee:

**Tania Elizabeth Calle Jiménez, Ph.D.**

Escuela Politécnica Nacional (EPN), Coordinator

**Sandra Patricia Sánchez Gordón, Ph.D.**

Escuela Politécnica Nacional (EPN), Opposing Member

**Sergio Luján Mora, Ph.D.**

Universidad de Alicante (UA), External Examiner

**Danilo Mauricio Pastor Ramirez, Ph.D.**

Escuela Superior Politécnica de Chimborazo ESPOCH, External Examiner

**Marco Eduardo Molina Bustamante, Ph.D.**

Escuela Politécnica Nacional (EPN), Internal Examiner

## DECLARATION

I declare under oath that I am the author of this work, which has not previously been presented for obtaining an academic degree or professional qualification. I also declare that I have consulted the bibliographic references included in this document.

Through this declaration, I transfer my intellectual property rights corresponding to this thesis to the Escuela Politécnica Nacional, as established by the Intellectual Property Law of Ecuador, its Regulations, and the current institutional norms.

I declare that this work is based on the following articles of my authorship (as main author or co-author) related to the title of this thesis:

### Journals:

- Martínez C., Moreno H., Hernández, M. (2023). Analysis of Intrusions into Computer Systems using Honeypots. *International Journal of Intelligent Systems and Applications in Engineering*, 11(6s), 461–472. Retrieved from <https://ijisae.org/index.php/IJISAE/article/view/2871>
- Martínez, C., Hernandez, M., Moreno, H. and Isa, R. (2024) "Classification of cognitive patterns of hackers using Machine Learning," in Volume (2024) of *Journal Computational Social Networks* ISSN21974314 has merged with the *Journal of Combinatorial Optimization*.

### Conferences:

- Martínez, C., Yoo, S., Moreno, H. (2018). " Analysis of Traditional Web Security Solutions and Proposal of a Web Attacks Cognitive Patterns Classifier Architecture "Technologies and Innovation. CITI 2018. *Communications in Computer and Information Science*, vol 883. Springer, Cham. [https://link.springer.com/chapter/10.1007/978-3-030-00940-3\\_14](https://link.springer.com/chapter/10.1007/978-3-030-00940-3_14)

- Martínez, C., Moreno, H. and Hernández, M. “The evolution from Traditional to Intelligent Web Security: Systematic Literature Review, In International Symposium on Networks, Computers and Communications (ISNCC), 2020, pp. 1–9. DOI:10.1109/ISNCC49221.2020.9297240

---

**(Carlos José Martínez Santander)**



## **CERTIFICATION**

I certify that Carlos José Martínez Santander has conducted his research under my supervision. To the best of my knowledge, the contributions of this work are novel.

---

**(Dr. Hugo Oswaldo Moreno Áviles)**

**ADVISOR**

---

**(Dra. Myriam Beatriz Hernández Álvarez)**

**CO ADVISOR**

## **DEDICATION**

I dedicate this work to my dear mother, María, who has been my greatest source of support and motivation throughout my life; thank you for your constant encouragement, sacrifices, and unconditional love.

To my loved ones, friends, and those who have been with me every step of the way, giving me love, encouragement, and understanding, as well as all the people who participated in this study and shared their time and knowledge with me, their contribution has been fundamental to the success of this project.

Finally, I dedicate this work to all those who relentlessly pursue knowledge and continuous improvement in their respective fields. May this contribution advance science and enhance the welfare of society. Thank you!

## **ACKNOWLEDGMENTS**

I would like to express my sincere gratitude to God for His constant blessings throughout my life.

I want to give a special thanks to the National Polytechnic School, the PhD Program in Computer Science, for allowing me to carry out my studies and especially to my supervisors, Dr. Myriam Hernández and Dr. Hugo Moreno, for their expert guidance, patience, and unconditional support throughout this process. Their advice and comments have been instrumental in shaping this work and in my academic and professional growth. My gratitude extends to my family, friends, and loved ones whose love, understanding, constant encouragement, and unconditional support have been my greatest strength.

To all of you, my sincere thanks.

# CONTENTS

<b>1</b>	<b>CONTEXTUALIZATION OF THE STUDY</b>	<b>1</b>
1.1	Introduction . . . . .	1
1.2	Justification . . . . .	2
1.3	Objectives . . . . .	3
1.4	Hypotheses . . . . .	4
<b>2</b>	<b>BACKGROUND</b>	<b>5</b>
2.1	Topics covered . . . . .	5
2.2	Computer Security Solutions . . . . .	5
2.3	Cybercriminal . . . . .	6
2.4	Artificial Intelligence . . . . .	6
2.5	Machine Learning . . . . .	6
2.6	Honeypots . . . . .	7
<b>3</b>	<b>RELATED WORK</b>	<b>9</b>
3.1	Introduction . . . . .	9
3.2	Area of Web Security . . . . .	9
3.3	The ELK Stack . . . . .	25
3.4	Honeypots . . . . .	26
3.5	Cybercriminal Personality . . . . .	35
<b>4</b>	<b>PATTERN CLASSIFICATION</b>	<b>47</b>
4.1	Materials and methods . . . . .	47
4.2	Web-Attacks Cognitive Patterns Classifier Architecture . . . . .	49
4.3	Analysis of the infrastructure T-Pot Honeypot . . . . .	51
4.4	Analysis of studies on the personality of hackers . . . . .	59
4.5	Application of the BFQ Survey . . . . .	63

<b>5</b>	<b>EVALUATION</b>	<b>71</b>
5.1	Dataset and Machine learning . . . . .	71
5.2	Discussion . . . . .	87
5.3	Contributions of this work . . . . .	89
<b>6</b>	<b>CONCLUSIONS AND FUTURE WORK</b>	<b>93</b>
6.1	Conclusions . . . . .	93
6.2	Future work . . . . .	96

## LIST OF FIGURES

3.1	Bibliographic Database . . . . .	11
3.2	The construction of subjective rating scales . . . . .	13
3.3	Evaluation of the quality of the Bibliography . . . . .	13
3.4	The predominance of articles and conferences 2014-2024 . . . . .	14
3.5	Web Security to Intelligent Web Security . . . . .	15
3.6	Attacks addressed to develop proposals or solutions in the field of web security	20
3.7	Current techniques and technologies applied to the improvement of web security	21
3.8	Shows the contribution of the researchers to the study problem. . . . .	22
3.9	Primary studies, year of publications Honeypots . . . . .	26
3.10	Primary studies, year of publications Cybercriminal Personality . . . . .	36
3.11	Word diagram of the motivation of cybercriminals . . . . .	41
4.1	WCPC architecture . . . . .	50
4.2	Server infrastructure . . . . .	52
4.3	Login T-Pot Honeypot . . . . .	53
4.4	Honeypot initialized ports . . . . .	54
4.5	IP of the ESPOCH homepage . . . . .	55
4.6	Attacks on the server implemented in the ESPOCH infrastructure . . . . .	55
4.7	The time zone from which the Attacks were launched . . . . .	56
4.8	Server with more registered attacks . . . . .	57
4.9	Countries from which attacks are made . . . . .	57
4.10	Database of Nicknames of Cybercriminals . . . . .	58
4.11	The Big Five Model (BFQ) . . . . .	62
4.12	Personality and attack data set . . . . .	64
4.13	Invite known elite hackers . . . . .	65
4.14	BFQ Hacker Elite Application . . . . .	66
4.15	BFQ Hackers Application . . . . .	67

4.16	Result about the type of hacker . . . . .	68
4.17	Findings about the type of elite hacker . . . . .	68
4.18	Findings on the personality of hackers . . . . .	69
4.19	Findings on the personality of elite hackers . . . . .	70
5.1	Classifier parameter contry . . . . .	72
5.2	Architecture of the Neuronal Network using a Sequential Model . . . . .	75
5.3	Confusion Matrix with the training data that the model already knew . . . . .	76
5.4	Confusion Matrix with training data not known to the model . . . . .	77
5.5	Density of model-predicted observations matched to actual observations . . . . .	78
5.6	Confusion Matrix of the model differentiates all attack classes correctly with known data . . . . .	79
5.7	Confusion Matrix of model differentiates all attack classes correctly with unknown data . . . . .	81
5.8	Classification decision tree . . . . .	82
5.9	Multi-correlation between hackers personality data . . . . .	83
5.10	Confusion Matrix with the personality test data known to the model . . . . .	84
5.11	Confusion Matrix with the personality test data that the model does not know . . . . .	86
5.12	Density of observations predicted by the model matched to the actual observations of the hacker patterns . . . . .	87

## LIST OF TABLES

2.1	A high and low interaction honeypot matching . . . . .	7
2.2	Description of honeypot types . . . . .	8
3.1	Findings in the field of web security . . . . .	19
3.2	Types, objectives, contribution, and Limitations of Honeypots . . . . .	34
3.3	Methods, Instruments, and Personality Classification of Hackers or Cyber-criminals . . . . .	40
3.4	Big Five Personality Traits . . . . .	45
4.1	Description of protocols . . . . .	52
4.2	Comparison of Big Five with other instruments . . . . .	60
5.1	Research parameters . . . . .	71
5.2	Structure of the confusion matrix . . . . .	73
5.3	Classification report of the training data the model already knew . . . . .	76
5.4	Classification report with training data not known to the model . . . . .	77
5.5	Classification report the model differentiates all attack classes correctly with known data . . . . .	79
5.6	Classification report of model differentiates all attack classes correctly with unknown data . . . . .	80
5.7	Type of parameters . . . . .	82
5.8	Classification report with the personality test data known to the model . . . . .	84
5.9	Classification report with the personality test data that the model does not know	85



## **ABSTRACT**

In the current digital era, cyberattacks pose significant threats to global security and privacy, presenting complex challenges that society must confront. Cybercriminals are employing increasingly sophisticated techniques to infiltrate systems and extract sensitive information. Understanding and anticipating the behavior of these attackers is essential for developing more agile and efficient defense strategies.

The main goal of this doctoral thesis is to develop a machine learning model to detect personality traits of potential cyber attackers using a combined approach of honeypot systems and survey research.

This research is methodologically structured into four key phases. Phase 1 involves information gathering, where honeypots are deployed to attract attackers and record their activities, providing valuable data on their tactics and behaviors. Phase 2 focuses on experimentation, where these collected data are used to identify initial patterns. Phase 3 employs a descriptive design, integrating survey responses from individuals involved in cyberattack activities to gather direct information about their personalities. Phase 4 involves the analysis of results, applying machine learning models to identify correlations between specific personality traits and cyberattack behaviors.

The results of this research contribute significantly to the field of cybersecurity. The developed models demonstrate the ability to predict certain personality traits based on behavior observed in honeypots and data collected through surveys. These findings expand our understanding of cyber attackers and provide a solid foundation for developing new cybersecurity defense strategies.

In conclusion, this work opens new opportunities for interdisciplinary research that integrates psychology and technology in the fight against cyberattacks.

**Keywords** - Machine Learning, T-Pot, Hacker, Cognitive Patterns, Attacker, Investigative Psychology

## PROLOGUE

In an increasingly globalized world dependent on technology, computer security has become a major challenge due to the growing number of cyber-attacks, their sophistication, and the evolving tactics of cyber-attackers. This situation creates a pressing need to understand not only the techniques and tools used by attackers but also their cognitive patterns and psychological profiles that drive their actions. This dissertation entitled: “LEARNING MODELS TO DETECT PERSONALITY TRAITS OF CYBERATTACKERS: A COMBINED APPROACH USING HONEYPOT AND SURVEYS” proposes an innovative and multi-faceted approach to detecting hacker personality traits. By combining honeypot techniques, designed to simulate vulnerabilities and attract potential attackers, with carefully designed psychological surveys administered within hacker communities, this research breaks traditional boundaries. It merges concepts, methodologies, and tools from computer science and psychology to provide a comprehensive and holistic perspective on the psychology of cyber-attacks. This multidisciplinary approach enhances our understanding of cybersecurity by incorporating insights into human motivation and criminal behavior in this digital world. This research takes us on an intellectual journey that encompasses both the technical aspects of computer security and the nuances of human psychology. The study explores various theories and methodological approaches to identify and classify relevant personality traits in the context of cyber-attacks. It covers a broad spectrum from a review of the field of study to the implementation of a T-Pot honeypot, qualitative data analysis from surveys, and the application of machine learning algorithms. This innovative research is valuable in encouraging other researchers to explore interdisciplinary approaches, highlighting the significant progress that can be made in computer security and psychology. It offers practical, updated ideas to improve the detection and prevention of cyber-attacks in an increasingly connected and vulnerable world. This work aspires to inspire future researchers to continue exploring the complexities at the intersection of technology and human behavior, ultimately contributing to effective and comprehensive solutions for protecting digital infrastructures.

# 1. CONTEXTUALIZATION OF THE STUDY

## Contents

---

1.1 Introduction . . . . .	1
1.2 Justification . . . . .	2
1.3 Objectives . . . . .	3
1.4 Hypotheses . . . . .	4

---

### 1.1. Introduction

Security is a crucial issue in various areas of computer sciences, including electronic devices [1], [2], networks, and software engineering, and internet applications are not exception [3]. According to the Internet Security Threat Report published by Symantec Corporation, web attacks were among the most frequent in 2016, with approximately 229,000 daily attacks. Information security is fundamental across computing areas [4]; new frameworks and technologies have been developed to ensure protection and counteract attacks on electronic devices, mobiles, networks, databases, cloud computing, and IoT, among others. Despite these technical measures, cybercriminals continue to grow in number. For decades, the Internet has suffered multiple attacks, and numerous defense schemes have been proposed to mitigate them [5].

A study by Symantec on vulnerability scanning in web servers showed that over three-quarters of websites worldwide have vulnerabilities, with 9% being critical. Cybercriminals often target technology and business websites due to their economic value. However, economic gain is not the only motivation for attackers. Many sites are hacked by hacktivist groups, often from countries in conflict, such as Syria and Pakistan, or by government-sponsored groups. A common technique used by these groups is defacement, which involves altering the visual appearance of a site. A list of servers hacked using this method can be found at [www.zone-h.org](http://www.zone-h.org) [6].

Both large and small businesses now offer web-based services to mitigate the problem of data exchange between the client and associated companies. Security challenges in web services are formidable [7]. Many companies process payments and manage confidential information transmitted over the web, making web security integral to daily life and access to various services. Critical online security failures can expose authentication credentials and confidential data stored on the page, leading to privacy violations [8].

In the last two years, several researchers have sought to stop attacks and strengthen computer security to the point of treating devices as extensions of the human body. These computational devices should mimic human decision-making and security features when faced with hacker threat. This innovative approach contrasts with traditional security measures, which rely on strong passwords, two-steps authentication, dialog boxes, and warnings. However, despite these measures, security remains fragile and vulnerable, often compromising comfort and productivity [9].

Given the growing cyber threat, new security solutions are continuously proposed. Cognitive security is one of the most significant trends. Cognitive systems help to improve security by monitoring threats globally and preparing for possible attacks. They also aid security analysts by providing human-centered tools such as advanced visualizations, interactive vulnerability analysis, and risk assessment [10], [11]. The three main pillars of cognitive security are: (1) organizing and understanding unstructured data and natural language text, (2) reasoning based on the ability to interpret and organize information, and (3) continuous learning from accumulated data and knowledge extracted from interactions.

## **1.2. Justification**

Cyberattacks have increased, making information and communication security a critical issue due to the constant increase in cyber threats and sophistication of attacks, malware, and denial of service (DDoS). Security measures include SSL/TLS certificates, authentication and authorization, SQL injection protection, web application firewall to prevent attacks such as brute force and Cross-Site Scripting XSS, and software updates to fix vulnerabilities. In this sense, cybercriminals have also evolved, finding new ways to exploit vulnerabilities in web servers or overcome existing security tools and technologies. Therefore, web security remains a major challenge, requiring continuous updates and awareness of new threats and proposed solutions.

Despite advances in all areas of computer security, significant gaps still exist related

to cognitive security. Cognitive security is an emerging approach that focuses on protecting computer systems by applying human cognitive processes. This involves systems that understand and mitigate vulnerabilities using machine learning algorithms trained for this purpose. One possible approach to cognitive security is to identify the personality profiles of hackers attacking systems. This information can be valuable for designing more effective security measures and outlining strategies to prevent or mitigate the consequences of cyber-attack. However, cybersecurity research using cognitive systems has not yet fully developed, representing a niche area. Developing learning models that detect the personality profiles of cyber attackers by combining two complementary methods is a novel approach.

The methodology of this thesis combines quantitative and qualitative analysis by collecting data from a honeypot and surveys. Honeypots simulate real systems to attract attackers, facilitating the collection of valuable information about their behavior and techniques. Surveys, on the other hand, provide insights into hackers' motivations and thought processes. This unique approach marks a significant contribution to the field of computer security, as no similar work currently exists.

### **1.3. Objectives**

**General Objective** To develop a machine learning model to detect personality traits of potential cyber attackers using a combined approach of honeypot systems and survey research.

#### **Specific Objectives**

- Analyze the state of the art of web security applied to servers and honeypots.
- Create profiles of cyber attackers based on their personality traits.
- Develop a dataset of attacks obtained through the implementation of honeypots.
- Apply personality surveys in hacker forums and to recognized hackers.
- Determine models generated using machine learning systems to detect personality traits of cyber attackers.
- Evaluate the results obtained and conclude how these results can be used to design more effective security measures and outline strategies to prevent or mitigate the consequences of cyberattacks.

The present research has the following results:

- State-of-the-art web security applied to servers and honeypots.
- Definition of cyber attacker profiles based on their personality traits.
- Data obtained through the use of honeypots.
- Results from applying surveys in hacker forums and to recognized hackers.
- Models generated using machine learning systems to detect personality traits of cyber attackers.
- Conclusions on how these results can be used to design more effective security measures and outline strategies to prevent or mitigate the consequences of cyber attacks.

## **1.4. Hypotheses**

Our hypotheses are:

- It is possible to create a dataset of cognitive patterns of hackers from the implementation of a honeypot and the application of surveys.
- The data of cognitive patterns left by hackers will provide enough information to classify them using a machine learning algorithm.

This document consists of six sections. The first section contains an introductory section on immersion in the research. Section 2 presents, as a problem, the fields on which the proposal is based. Section 3 is the state of the art of the fields on which the proposal is based. Section 4 describes this work's materials, methods, and techniques. Section 5 presents the generated dataset and the modeling using the selected parameters as input for the machine learning algorithms to classify the cognitive patterns of hackers with their respective results. Finally, Section 6 presents the conclusions of the research.

## 2. BACKGROUND

### Contents

---

2.1	Topics covered . . . . .	5
2.2	Computer Security Solutions . . . . .	5
2.3	Cybercriminal . . . . .	6
2.4	Artificial Intelligence . . . . .	6
2.5	Machine Learning . . . . .	6
2.6	Honeypots . . . . .	7

---

### 2.1. Topics covered

This section provides an overview of the topics covered in the proposal, including computer attacks, types of hackers, honeypots, computer security, and machine learning, and others. It also explains key concepts and technologies used in this document.

### 2.2. Computer Security Solutions

Computer security solutions have evolved rapidly over time. Around the year 2000, the focus was primarily on detecting, analyzing, and eliminating malicious code using specialized software, such as antivirus programs, to provide endpoint protection [10]. By 2005, the emphasis shifted to security solutions focused for network devices that could detect anomalies in network traffic and alert security personnel to unusual activity. During this period, the term “security intelligence” gained popularity [11]; this technology involved analyzing large volumes of real-time data to identify potential threats. By 2010, security technologies had become more advanced, encompassing not only network traffic but also server logs, and various devices.

This shift marked a transition from reactive to proactive security [12]. Since 2015, the concept of Cognitive Security has emerged. Building on the principles of security intelligence,

Cognitive Security leverages artificial intelligence to simulate the human thought process. These systems can analyze both structured and unstructured data, understanding behavior and meaning, and identify vulnerabilities by connecting data points, detecting anomalies, and analyzing vast amounts of events to enhance the knowledge database [13].

### **2.3. Cybercriminal**

Cybercriminals gain unauthorized access to computer systems to commit fraud and abuse. Numerous studies have focused on understanding the motivations behind these disturbances. Research indicates that many attackers are key employees within their own companies, engaging in such activities for professional or economic recognition, and in some cases, for personal enjoyment or revenge [12].

### **2.4. Artificial Intelligence**

The goal of artificial intelligence (AI) is to replicate human reasoning in machines. This is achieved through two main activities: (1) imitating human reasoning methods, which involves psychological experiments and cognitive science, and (2) replicating brain functionality, which is supported by neurobiology [14].

### **2.5. Machine Learning**

Machine learning is a subfield of artificial intelligence focused on optimizing information processing through theories from information theory, statistics, and cognitive science. This technique enables computers to automatically discover, recognize patterns, and make predictions. It also uncovers underlying relationships within large datasets. Machine Learning has evolved significantly with the increasing volume and complexity of data [14]–[18].

Machine learning has been applied to various problems across different domains, including search engines, automatic control systems, recognition systems, and data mining. This field is generally divided into two categories:

- Supervised learning, which involves learning from labeled data to predict desired inputs and outputs.
- Unsupervised learning, which involves learning from unlabeled data, only inputs without desired goals and reinforcement [16].



## 2.6. Honeypots

Honeypot is a security tool designed to be explored, attacked, and compromised, with the aim of detecting and identifying malicious activities. Honeypots is valuable for understanding various types of attacks and for gathering information about security issues. By appearing as a real server with valuable information, a honeypot attracts attackers, who invest their efforts in exploiting it [19].

Honeypots can be categorized as high-interaction and low-interaction [20], [21], High-interaction honeypots are more complex and expensive to implement but provide a comprehensive and realistic view of attackers' techniques and patterns. In contrast, low-interaction honeypots are simpler, cheaper, and less detailed, often being more easily detected by attackers [22], [23]. The choice of using a high-interaction or low-interaction honeypot depends on the specifics needs [24], goals, and resources of the organization or research project. Table 2.1. shows a comparison between these two technologies.

<b>Low Interaction Honeypot</b>	<b>High Interaction Honeypot</b>
The attacker can detect it easily and quickly, which limits its efficiency.	Its detection is more difficult for the attacker, which translates into effectiveness in capturing and recording attacks and threats.
It requires the minimum investment of resources for its application and maintenance,	It requires a greater investment of resources and time for its implementation and maintenance.
It does not generate detailed data. That is, it provides an overview of the captured attacks.	Detailed and real data on attack methods and the tools used by the attackers.
It simulates a vulnerable system without allowing the attacker to develop his skills.	The attacker interacts with the system, leaving a record of the tactics and techniques used.
It can be deployed to existing operating systems and applications.	It requires a complete operating system environment and applications to be deployed.

Table 2.1: A high and low interaction honeypot matching

T-Pot Honeypot is a sophisticated honeypot system known for its ability to emulate multiple services and operating systems. It stands out from other solutions due to its integration

with Security Information and Event Management (SIEM) systems, its user-friendly installer, and complete documentation. Table 2.2 below provides a comparison of honeypots based on insights from various research studies.

<b>Honeypot</b>	<b>Interaction</b>	<b>Customiza tion</b>	<b>Integration</b>	<b>Virtualiza tion</b>	<b>Difficulty of use</b>	<b>Integration with SIEM</b>
T-Pot	High/Low	High	Yes	Yes	Medium	Yes
Honeypot Honeyd	High	High	No	Yes	High	No
Dioneda	High	Medium	Yes	Yes	Medium	No
KFSensor	High	High	Yes	No	High	Yes
Glastopf	Low	Medium	No	Yes	Medium	No
Amun	Low	High	Yes	Yes	High	No

Table 2.2: Description of honeypot types

## 3. RELATED WORK

### Contents

---

3.1 Introduction . . . . .	9
3.2 Area of Web Security . . . . .	9
3.3 The ELK Stack . . . . .	25
3.4 Honeypots . . . . .	26
3.5 Cybercriminal Personality . . . . .	35

---

### 3.1. Introduction

This section reviews the latest research, theories, and advancements in the fields relevant to this project. This critical analysis of the literature highlights knowledge gaps and opportunities for significant contributions to the field. Note that the doctoral project commenced around 2017, and therefore, the information covers various periods up to the present, which may explain the absence of some recent research.

### 3.2. Area of Web Security

This study follows the systematic literature guidelines in software engineering established by Kitchenham [25]. The research questions guiding this study are:

- RQ1. What are the primary attacks addressed to develop proposals or solutions in web security?
- RQ2. What are the comprehensive solutions proposed in the field of web security?
- RQ3. What are the current techniques and technologies applied to enhance web security?
- RQ4. What are the contributions of existing research in the field of web security?

- RQ5. What are the limitations of current research?
- RQ6. What are the proposed directions for future work in web security?

RQ1 seeks to identify the primary attacks that researchers have addressed in web security, focusing on the most common attacks over the last five years. At the same time, this question helps to determine whether there are definite or partial solutions available. A systematic review of literature from 2016 [26] was identified during the search process. However, this study lacked the rigor required for a comprehensive SLR and is outdated given the rapid technological advancements and sophisticated cyber-attacks in recent years. This underscores the need for the current SLR.

RQ2, RQ3, and RQ4 analyze existing solutions or those proposed by researchers to determine if there are comprehensive solutions for information security, including prevention, detection, and practical response to attacks. Effective solutions should enable the prevention of attacks, provide countermeasures for active threats, and significantly reduce vulnerabilities in web servers. RQ5 outlines the limitations of current solutions identified in primary studies, while RQ6 discusses proposed future work based on the results and experiences gained for this field of web security.

### **Search Process**

The search process was carried out on the primary high-level databases, using search strings designed using the inclusion and exclusion criteria, according to the format of each search engine, for example for Scopus: "TITLE-ABS-KEY (WEB AND ATTACKS OR WEB AND SECURITY OR INSTRUCTION AND DETECTION OR WEB AND APPLICATION AND SECURITY OR WEB AND DEFENSES) AND (LIMIT-TO (PUBYEAR, 2019) OR LIMIT-TO (PUBYEAR, 2018) OR LIMIT-TO (PUBYEAR, 2017) OR LIMIT-TO (PUBYEAR, 2016) OR LIMIT-TO (PUBYEAR, 2015) TO (PUBYEAR, 2014)". For Web of Science: "TI=(WEB AND ATTACKS OR WEB AND SECURITY OR INSTRUCTION AND DETECTION OR WEB AND APPLICATION AND SECURITY OR WEB AND DEFENSES) REFINATE TO: YEAR OF PUBLICATION: (2019 OR 2018 OR 2017 OR 2016 OR 2015 OR 2014) AND WEB CATEGORIES OF SCIENCE: (COMPUTER SCIENCE INFORMATION SYSTEMS OR TELECOMMUNICATIONS)". Keywords, combined with logical operators, were used in the ScienceDirect search engine, as in ACM Digital Library; "WEB AND ATTACKS OR WEB AND SECURITY OR INSTRUCTION AND DETECTION OR WEB AND APPLICATION AND SECURITY OR WEB AND DEFENSES". Finally, in IEEE Xplore Digital Library: "(("INDEX TERMS": WEB ATTACKS) AND DETECTION) AND DEFENSE) FILTERS APPLIED: 2014-2018", Fig-

ure 3.1 shows the results of the queries performed.






Source	Reference
	385
	742
	282
	103
	36

Figure 3.1: Bibliographic Database

#### Inclusion and Exclusion Criteria

- Research published between January 2014 and 2024 in journals related to Computer Science, Computer Science Security, or Electronics.
- Studies presenting solutions to web security issues.
- Research that clearly outlines the methodology used to obtain the results.
- Researches utilizing methodologies and technologies for detecting vulnerabilities in web services.

#### The exclusion criteria were:

- Bibliographic review articles.
- Comparative studies or research that present only theoretical contributions without empirical results.

- Duplicated studies or those unrelated to web service attacks, including cross-site scripting (XSS), Denial of Service (DoS) attacks, SQL injection, website defacement, and XPath injection.

### **Quality assessment**

- Each study was evaluated using the criteria outlined by Kitchenham in the RSL guide [25]. The qualitative assessment involved the following questions:
  - Are the findings credible?
  - Are the findings significant?
  - Has the research contributed to extending knowledge or understanding?
  - Does the evaluation effectively address the original research objectives?
  - Is the scope for broader inference well-explained?
  - Are the study design, case selection, and documentation robust?
  - Was data collection conducted appropriately?
  - Is the approach and formulation of the analysis clearly conveyed?
  - Are the connections between data, interpretation, and conclusions clear and logical?
  - Is the information presented consistently and clearly?

Note: Each item was scored using a Likert scale [27], The total average and absolute standard deviation were used to define the relevance of the studies for inclusion. Out of the 58 preselected items, ten were excluded for scoring less than or equal to 3.10, as detailed in Figures 3.2 and 3.3.

### **Data Collection**

After selecting the primary studies, the following data was collected from each article:

- Code, Title, Year of publication.
- Paper published in a magazine or a conference.
- Abstract, Conclusions, limitations, and future works.
- Area of focus, Type of attack, Technique, and Contribution.

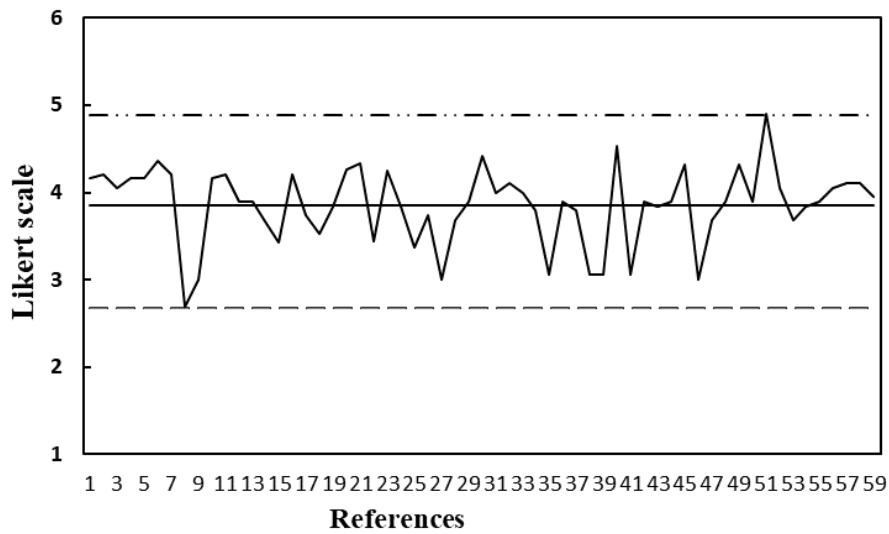


Figure 3.2: The construction of subjective rating scales

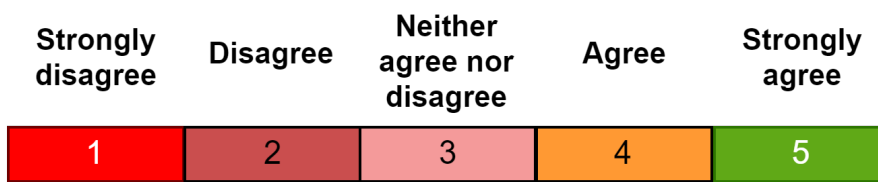


Figure 3.3: Evaluation of the quality of the Bibliography

### Data Analysis

The data was analyzed using criteria related to the research questions (RQ) and tabulated as follows:

- Types of attacks addressed: DOS attack, Website Defacement, XPath injection, Cross-site scripting (XSS), SQL injection, XML injection, Clickjacking, Spoofing (Directed to RQ1).
- Techniques used researchers: Dynamic Analysis, Static Analysis, Model- Based, and Secure Programming (Directed to RQ2 and RQ3).
- Techniques for analyzing the environment or data (directed to RQ3).
- Areas of focus: attack detection, attack prevention Combination, and vulnerability detection (Directed to RQ2 and RQ3).
- Solutions proposed: Algorithm, Application, Architecture, Frame- work, IDS, IPS, Method,

Model, System, Tool, and Security politics (Directed to RQ4).

- Conclusions and future work: contributions and proposed future directions (Directed to RQ5 and RQ6).

The studies primarily published in journals and conferences, with a noticeable increase in publication in recent years, as shown in Figure 3.4. A predominance of articles published in conferences is observed; likely due to faster publications timelines compared to journals.

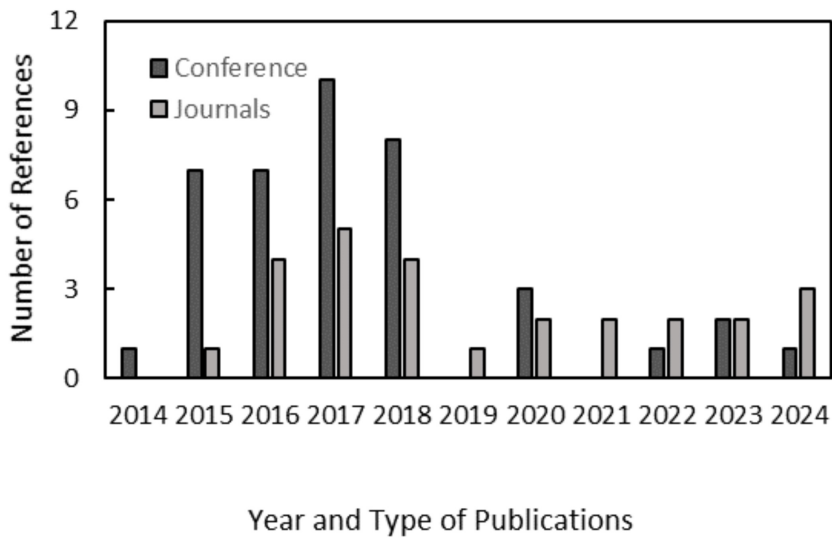


Figure 3.4: The predominance of articles and conferences 2014-2024

Regarding the use of machine learning algorithms in [28], a meta-classifier, a combination of predictor trees known as Random Forests, is used. In [29], regression metrics are used for data analysis, complemented by Euclidean distance analysis and Chebyshev's inequality theorem to compute normality. While in [30] and [31], they often use Naive Bayes supervised algorithm and J48 Tree frequently; this algorithm is used to generate decision trees using the Weka software. In [32] and [33], they used ML techniques such as Decision Tree, Rule-based, Support Vector Machine, Neural Network, and Random Forest algorithms using Weka tools. In [25], they use the decision tree recursively; all these results are shown in Figure 3.5.



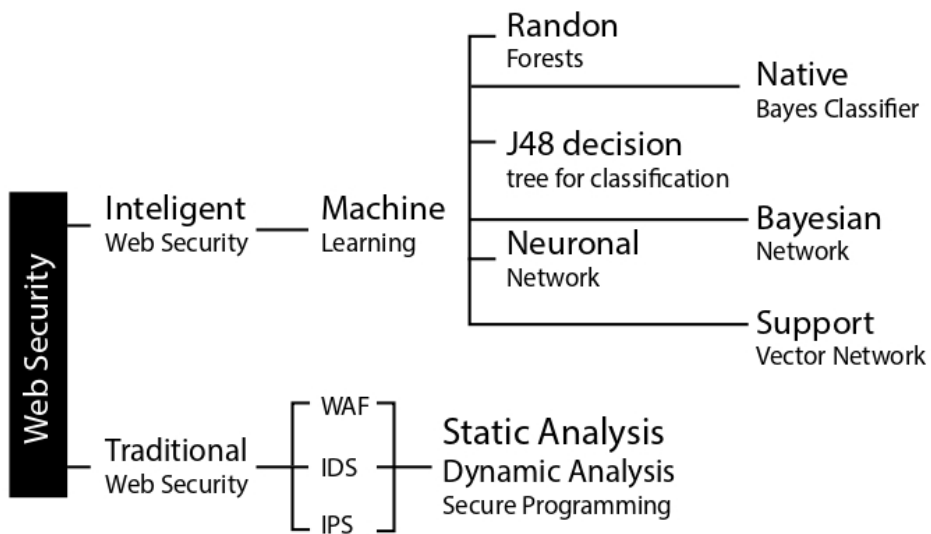


Figure 3.5: Web Security to Intelligent Web Security

Table 3.1 summarizes the key conclusions drawn from various research studies in the field of Internet security. This table provides an overview of significant findings, including the techniques used, their contributions, and the main types of attacks addressed. It highlights the effectiveness of various methods and tools, such as honeypots, in detecting and mitigating threats within the digital environment.

<b>Ref</b>	<b>Area of Focus</b>	<b>Techniques</b>	<b>Contribution</b>	<b>Attacks</b>
[29]	Attack detection	Static Analysis	Algorithm	DOS attack
[30]	Attack prevention	Secure Programming	Application	SQL injection
[31]	Attack detection	Dynamic Analysis	IDS	Several attacks
[32]	Vulnerability detection	Static Analysis	Method	Several attacks
[33]	Attack detection	Static Analysis	IDS	Several attacks
[34]	Attack prevention	Dynamic Analysis	Architecture	Cross-site scripting (XSS)
[28]	Attack detection	Model Based	Model	Website Defacement
[35]	Attack detection	Model Based	System	XPath injection
[36]	Attack detection	Model Based	Tool	Website Defacement
[37]	Attack detection	Dynamic Analysis	Method	Cross-site scripting (XSS)
[38]	Attack detection	Dynamic Analysis	System	Cross-site scripting (XSS)
[39]	Combination	Dynamic Analysis	System	Other attacks
[40]	Vulnerability detection	Model Based	Model	Cross-site scripting (XSS)
[41]	Vulnerability detection	Dynamic Analysis	Tool	Cross-site scripting (XSS)
[42]	Attack detection	Model Based	Framework	Several attacks
[43]	Attack prevention	Model Based	Algorithm	Cross-site scripting (XSS)
[44]	Combination	Model Based	Framework	Several attacks
[45]	Attack detection	Dynamic Analysis	Method	Several attacks
[46]	Combination	Model Based	Algorithm	SQL injection
[47]	Attack detection	Static Analysis	IDS	Several attacks

[48]	Combination	Dynamic Analysis	Framework	Other attacks
[49]	Combination	Dynamic Analysis	System	SQL injection
[50]	Combination	Dynamic Analysis	System	Several attacks
[51]	Combination	Dynamic Analysis	System	Cross-site scripting (XSS)
[52]	Attack detection	Static Analysis	Method	SQL injection
[53]	Attack detection	Static Analysis	Method	DOS attack
[54]	Vulnerability detection	Static Analysis	Application	Several attacks
[55]	Attack prevention	Secure Programming	Model	Several attacks
[56]	Attack prevention	Dynamic Analysis	Algorithm	DOS attack
[57]	Vulnerability detection	Dynamic Analysis	Algorithm	DOS attack
[58]	Combination	Static Analysis	Method	DOS attack
[59]	Attack detection	Dynamic Analysis	IPS	DOS attack
[60]	Attack prevention	Secure Programming	Framework	SQL injection
[61]	Combination	Dynamic Analysis	System	SQL injection
[62]	Attack detection	Dynamic Analysis	Tool	DOS attack
[63]	Combination	Other	Architecture	Several attacks
[64]	Attack detection	Dynamic Analysis	IDS	Several attacks
[65]	Combination	Dynamic Analysis	Framework	Several attacks
[66]	Attack detection	Dynamic Analysis	System	Several attacks
[66]	Combination	Other	Tool	Other attacks
[67]	Attack detection	Dynamic Analysis	IDS	Several attacks

[68]	Attack prevention	Secure Programming	System	Several attacks
[69]	Attack prevention	Secure Programming	Security politics	Several attacks
[70]	Attack detection	Dynamic Analysis	Architecture	SQL injection
[71]	Combination	Model Based	Tool	Website Defacement
[72]	Combination	Dynamic Analysis	IDS	Several attacks
[73]	Combination	Dynamic Analysis	Architecture	Cross-site scripting (XSS)
[74]	Attack detection	Static Analysis	Algorithm	DOS attack
[75]	Attack detection	Fuzzy AHP Fuzzy TOPSIS	Application	Attacks SQL injection, cross scripting
[76]	Attack detection	Static Analysis	Algorithm	Cross-site scripting (XSS)
[77]	Attack evaluate and compare	Logic Score of Preferences	Framework	DOS attack
[78]	Attack detection	Static, dynamic and interactive analysis	Tool	DOS attack
[79]	Attack detection	Machine learning	Architecture	Attack SQL injection, CRLF injection, XSS
[80]	Attacks prevent and detect	SQLIA	Model	SQL Injection Attacks (SQLIA)
[81]	Vulnerability detection	Analysis dynamic	Model	DOS Attack
[82]	Attack detect	Analysis dynamic	Model	Attacks SQL Injection, cross site scripting

[83]	Combination	Analysis dynamic	Method and tools	Attacks SQL Injection
[84]	Attack detect	Analysis dynamic	Framework	DDoS Attack
[85]	Attack detection and prevention	Logistic Regression, SVM, TFIDF vectorizer, XGBoost	Machine learning model	SQL injection attack
[86]	Vulnerability detection	Dynamic and interactive analysis	Methodology	SQL injection attack

---

Table 3.1: Findings in the field of web security

Figure 3.6 illustrates that 43.18% of references focus on solutions designed to address or counteract multiple concurrent attacks, such as preventing SQL injection and blocking a DOS attack [29], [28]–[47]–[49]. The 18.18% of the research examines solutions for SQL injection, which account for 11.36% of the studies [31], [41], [50], [53], [54] [55]–[62], [33] [63]–[65].

These solutions range from secure programming techniques to advanced models aimed at detecting and preventing this vulnerability, largely due to its prevalence. One major cause of SQL injection vulnerabilities is the use of insecure coding practices by novice programmers, who may inadvertently include SQL code in their applications. Finally, the 6.82% and 2.27% of the research address Website Defacement and Xpath injection attacks, respectively [30], [66]–[68].

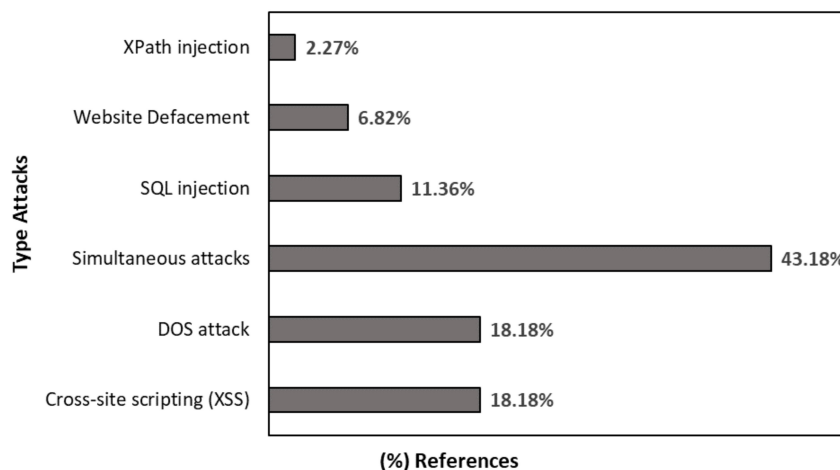


Figure 3.6: Attacks addressed to develop proposals or solutions in the field of web security

The studies proposing comprehensive solutions include [32], [39], [45], [47], [69]. These studies, published in conferences, focus on combined approaches that prevent and detect both attacks and vulnerabilities at the webserver level. For example, in [34], the author designed, implemented and evaluated a framework over seven months, yielding acceptable results. However, it also exhibited a high false negative rate, indicating it may not be reliable as a standalone solution.

SQL injection attacks [65] pose a significant threat to the security of financial and critical data. The proposed solution involves using of rule-based algorithms combined with decision trees, closely integrated with neural networks, to improve data classification efficiency. An Architecture combining record acquisition technologies and semantic web attempts to

address current security-analysis challenges [70]. However, it remains a prototype, and its effectiveness in real-world environments is unknown.

Hong et al. [71] developed a tool that complements a vulnerability scanner still under development, so its full potential is not yet known. Thus far, the tool addresses the generation of 602 simple and sophisticated attacks. Other research develops an IDS [87] based on signatures for web services, aiming to create new attacks trademarks, though its effectiveness is still being tested.

Commercial solutions have made significant progress in preventing, correcting, and treating many vulnerabilities. However, these technologies and algorithms are not open to the public or researchers and involve substantial costs, making them inaccessible to small and medium-sized enterprises. The primary techniques used in web security research are summarized in Figure 3.7.

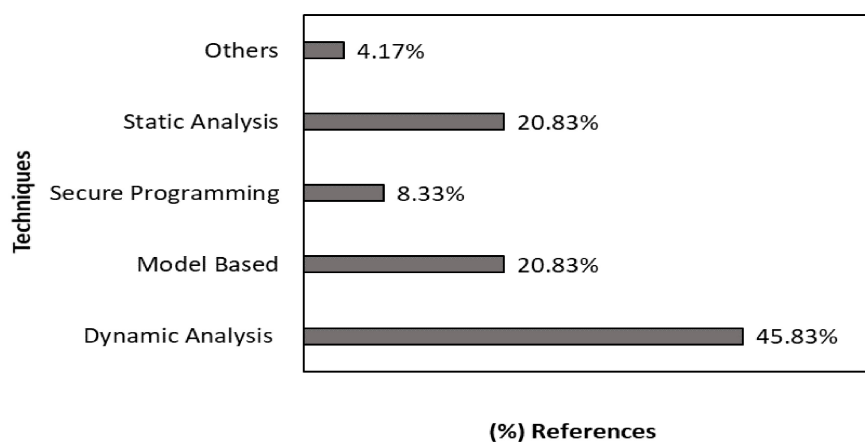


Figure 3.7: Current techniques and technologies applied to the improvement of web security

Annexing the data in Figure 3.7, it can be seen that most of the research employs Dynamic Analysis techniques[31], [33], [28], [36], [39], [41], [42], [61],[51], [52], [54], [72]–[74], [88], [89] accounting for 40.91%. This approach has significantly advanced the detection of attacks by analyzing unusual data patterns and inconsistencies in input and output data based on execution time.

In contrast, 18.18% of the researchers propose solutions using Static Analysis techniques. While these methods were once highly effective, their efficacy has diminished due to advancements in techniques such as Machine Learning, Deception Evaluation, and Pattern Classification [29], [40], [43], [71], [90], [91].

Approximately 15.91% of the solutions are model-based [28], [30], [50], [63], [67], [92].

Meanwhile, 13.64% of the solutions focus on statistical correlations, semantic approaches, and user-level applications (e.g., browser-based applications) to prevent attacks [45], [53], [60], [72], [93]. Lastly, 11.36% employ secure programming techniques designed to prevent attacks during the design and testing phases of application development, thereby avoiding vulnerabilities in production [37], [64], [65], [69].

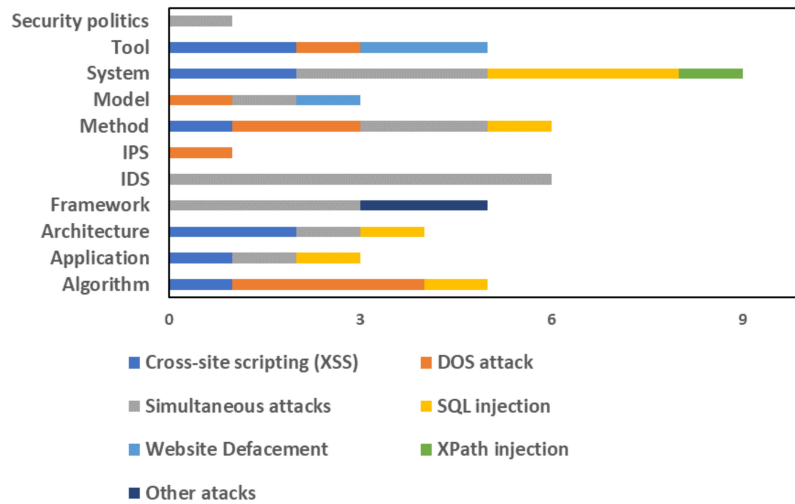


Figure 3.8: Shows the contribution of the researchers to the study problem.

Figure 3.8 illustrates the diverse contributions from researchers, including nine system designs, six models, six Intrusion Prevention Systems (IPS), five tools, five algorithms, four frameworks, three architectures, and one study proposing security policies. Many of these contributions have not been thoroughly tested, making it difficult to assess their limitations and the potential of the proposed tools. Sachdeva et al. [94] have as a limitation not being able to validate the metrics of detection of real traffic. In [65], the limitations go in the sense that the scanners of web applications must update for the automatic detection of vulnerabilities SPARQL / SPARUL.

In [51], the false positive rate remains still very high, indicating that vulnerability detection mechanisms should be improved. In [42], the communication functionality only implies delay, and hardware with enough resources is needed, implying high acquisition and maintenance costs. Marashdih [50] focuses solely on vulnerabilities at the PHP level, suggesting that future research should also consider other programming languages given the increasing number of applications not developed with PHP. Mewara et al. [95] presents a great approach but does not reduce CPU performance and employs many phases to determine the existence of vulnerabilities or attacks. The research [69] does not confirm the effectiveness



of the attack detection technique, thus its limitations remain unidentified. Hofstede et al. [68], mention that their detection method is not very resistant to the evasion technique. In [88], the solution's functionality is not yet thoroughly evaluated.

On the other hand, in [89], it does not work for all databases and should be automatically updated with new injection attacks information. Rao et al. [48] found that while ModSecurity's firewall filters code to detect and prevent threats, it would quickly succumb to large-scale attacks. The research in [46] cannot handle encrypted information, and the distinction between attack and defense is insufficient. Shahriar [47] proposes an intelligent system based on sensors, but its effectiveness is not guaranteed.

Not all the researchers present future works, but those who offer improvement to their systems:

Zhao et al. [92] aim to improve the experimental conditions of their environment, examine the complexity of their defense technique more thoroughly, and increase the number of nodes. They also plan to enhance frame detection capabilities and add indicators to measure attack intensity, such as traffic volume and the number of packets. Ashar [65] proposes studying a second class of attack, "stored SPARQL injection," to analyze vulnerabilities in libraries that provide SQL / SPARQL interoperability.

Kamrudin et al. [29] suggest transforming successful results into signatures to be stored in the blacklist database for future identification.

Research [41], plans to implement their system in a real environment and measure its efficiency and accuracy. In [67], they develop deep learning techniques that can explain so that the classifier bases the reasons behind its classification decisions for each incoming query so that more previously unknown attack patterns are discovered; in the research [87] they intend to have a mobile application that allows the same functionalities as the WDIMT website.

Zhou [31] proposes to apply this method in many more data sets and practical scenarios, as well as to integrate the proposed method and use its results in a real system assessing security risks on the web.

In [51], the plan to improve the false-negative rate includes implementing a mechanism for detecting of XSS vulnerabilities based on the DOM. Additionally, there is an intention to extend the use of the proposed framework to discover other vulnerabilities such as Phishing attacks and Click-Jacking. The final step involves evaluating the detection system's discov-

ery capacity to identify vulnerabilities in more web applications.

Marashdih [50] suggests that future work should focus on developing comprehensive solutions for both the detection and elimination of XSS vulnerabilities. Similarly, Perera et al.[49] propose implementing neural networks to enhance detection processes across various applications. Relatedly, the study in [68] suggests that future methods should be made more resistant to evasion techniques.

Sultana [70] emphasizes the importance of applying feature selection measures to improve the accuracy of classifiers, while Gupto [88] proposes evaluating the attack vector detection capability of his work in real-world web applications. Hou [89] also highlights the need to examine new injection possibilities and other vulnerabilities, particularly in NoSQL databases and emerging platforms.

Ross [32] suggests that to improve the performance and accuracy of learning techniques, data collection from the web application to the browser should be included. Similarly, the study in [34] suggests complementing existing DoS request methods with an extended detection approach and advanced detection expression.

Ishikawa [43] considers that his proposal could benefit from two key improvements: refining the experiment and adjusting the decoy parameter. Laskar [72] recommends adding additional features to identify different types of DDoS attacks. In another approach, the research in research [33] plans to implement an automatic training system for new genuine or attack sequences detected outside the original dataset.

The study in [45] focuses on interpreting individual events and semantic techniques to establish relationships between facts and link them to security knowledge. Souissi [46] suggests that future systems could include managing encrypted information and metrics to improve the alignment between attack and defense.

Additionally, the work in [47] considers applying multiple crossing points, developing physical conditioning functions with WSDL information, and conducting large-scale evaluations of web services, including an anomaly-based IDS. Zhang et al. [35] recommend using control systems theory and game theory to improve the design of cyber defense systems, while Alsaleh et al. [36] propose integrating data from other security logs into the PHPIDS visual-

ization extension.

Finally, in [64], the aim is to limit the load on computing power and network bandwidth caused by expanded test coverage by prioritizing test combinations with the highest chances of discovering vulnerabilities. On the other hand, [66] considers allowing commands to run on a Windows machine within the WDMIT, in the study [39] it can include an application of multiple cross points.

### **3.3. The ELK Stack**

For [51], the ELK stack, which combines open-source tools Elasticsearch, Logstash, and Kibana, operates in a virtual environment to provide a comprehensive approach to data analysis and consolidation. Each component plays a specific role: Elasticsearch is an Apache-based search engine, Logstash collects necessary logs and sends them to Elasticsearch, converting them into JSON format, and Kibana visualizes this data in tables, figures, or maps. Due to these advantages, ELK is used to build robust security log analysis systems for companies, offering a low-cost alternative to commercial products and helping startups avoid the lengthy process of creating less effective log systems. ELK has a proven record in log time recording and offers various visualization tools to assist security administrators, making it a powerful element in IT security analysis.

In the study of [52], the authors propose integrating the ELK stack with SIEM and PACK as an alternative security solution. By combining these components, companies handling large amounts of data can achieve robust security. This integration leverages each component's features, such as security, machine learning, and alerts, focusing on protecting privacy and data processed on certain platforms. This approach urges stakeholders to prioritize data security and seek effective protection alternatives.

In [53], the authors highlight ELK's efficiency in log analysis and its user-friendly tools. The log systems supported by ELK are designed to handle large data volumes and facilitate process tracking and calculation via an interactive interface. As an open-source solution, it offers extensive capabilities for log analysis. Elasticsearch serves as the indexing, storage, and retrieval mechanism, Logstash acts as the input receiver, parser, and output writer, and Kibana manages visualization through the dashboard. The implementation of ELK in this study effectively targeted web users for logging purposes.

Chen et al. in [54] discuss the challenges in Docker cluster log collection, including low efficiency, weak implementation, and poor stability. They propose adopting the ELK stack, Filebeat, and Kafka to design a Docker container log collector and analyzer system. This system, supported by ELK, enables rapid deployment for real-time log collection, filtering, and data visualization and analysis, significantly improving staff efficiency. The solution demonstrates good real-time performance, stability, and high availability.

### 3.4. Honeypots

It is important to note that the scientific production on the topic of honeypots has been frequently addressed by researchers in the last decade because it is a technique widely used by information security managers to know the types of attacks a cyber attacker uses. In Figure 3.9, you can find publications from 2018 to 2023, where you can observe high and low peaks in certain years of scientific production on this topic, a slight increase in publications in 2020, and a similar number of publications in the other years.

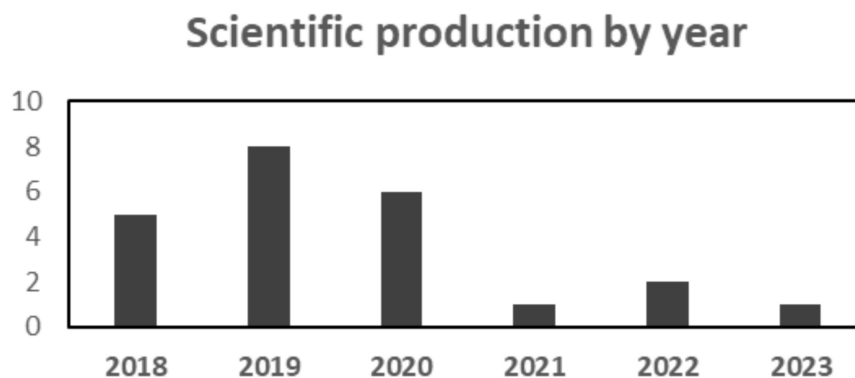


Figure 3.9: Primary studies, year of publications Honeypots

Table 3.2 presents the research works identified, providing details on the types of honeypots used, their main objectives, contributions, limitations, and whether they employ machine learning techniques for detecting or intruding on cyberattacks. Notably, most of the studies utilize high-interaction T-POT honeypots due to their effectiveness in simulating real services.

Ref	Year	Title	Honey type	Purpose	Machine learning	Contribution	Limitation
[96]	2018	Dynamic Honey-pot Configuration for Intrusion Detection		To Identify unauthorized access and network intruders.	N/A	Dynamic Honey-pot Configuration.	N/A
[97]	2018	Investigation of modern attacks using proxy honey-pot	Proxy honey-pot	Modern Attack Research	N/A	Improved detection and response to network intrusions	need to complement the use of proxy honey-pots.
[98]	2018	Hybrid System Between Anomaly-Based Detection System and Honey-pot to Detect Zero-Day Attack	Hybrid System	Zero-Day Attack Detection	N/A	Improvement to detection of zero-day attacks	Complexity of Implementation, Maintenance and Detection
[99]	2018	An SSH Honey-pot Architecture Using Port Knocking and Intrusion Detection System	High-interaction	Increasing Security and Intrusion Detection	N/A	Intrusion detection in SSH services	Complexity of implementation and resources required

[100]	2018	Honeypots That Bite Back: A Fuzzy Technique for Identifying and Inhibiting Fingerprinting Attacks on Low Interaction Honeypots	Low-interaction	Identification and Mitigation of Fingerprinting Attacks.	N/A	Fuzzy technique for improved fingerprinting attack detection and mitigation	Complexity, resources and scalability.
[101]	2019	Probabilistic Estimation of Honeypot Detection on the Internet of Things Environment	Medium-interaction	IoT Intrusion Detection and Mitigation	N/A	Probabilistic Methodology and the Implementation of Automated Software	Sophisticated Attacker Detectability
[102]	2019	Multiplatform Honeypot for the Generation of Cyber Threat Intelligence	Low interaction	Cyber Threat Intelligence Generation	supervised and unsupervised	Multi-Platform Support, Advanced Threat Analysis and Intelligence Reporting	implementation and possible detectability by adversaries.

[103]	2019	A honeypot with machine learning-based Detection framework for defending IoT based botnet DDoS attacks	Combined	Defense against IoT DDoS Attacks	Supervised and Unsupervised	Advanced Detection Framework Real-Time Analysis	Performance and Production Environments	in
[104]	2019	A Novel and Interactive Industrial Control System Honeypot for Critical Smart Grid Infrastructure	Combined	Critical Infrastructure Defense and Security	Data Analysis	Interactivity and Realism, Threat Intelligence Generation	Interaction Realism	
[24]	2019	Data Analytics Layer For high-interaction Honeypots	High-interaction	Advanced Data Capabilities	Data Analysis	Data Analysis and Threat Intelligence Enhancement Layer	Complexity and Scalability	and
[105]	2019	HoneyDOC: An Efficient Honeypot Architecture Enabling All-Round Design	HoneyDOC	Simulation of Attractive Documents	N/A	Integral Design and Data Capture Efficiency	Sophisticated Attacker Detection	

[106]	2019	The Security of Heterogeneous Systems based on Cluster High-interaction Hybrid Honeypot	High-interaction	Improvement of Computer Security.	N/A	Security System Design	Detectability by Sophisticated Attackers
[107]	2019	Automatic identification of honeypot server using machine learning techniques	High-interaction	Honeypot Server Detection	Techniques	Development of Machine Learning Techniques	Attacker Detectability
[108]	2020	Using Global Honey-pot Networks to Detect Targeted ICS Attacks	High-interaction	Detecting and profiling targeted attacks against ICS	N/A	Detection of malicious attacks such as DoS and protocol tampering	Honeypots updated
[109]	2020	Implementation of an insider threat detection system using honeypot-based sensors and threat analytics	Encryption	Detect insider threats and misuse of information		Technique to improve insider trading detection, reducing false positives	Implementation in real scenarios



[110]	2020	HONEYDOS: a hybrid approach using data mining and Honeybot to counter denial-of-service attacks and malicious packets	Low-interaction	Prevent denial-of-service attacks through honeypots and data mining.	Data mining	Honeydos combining support machines and data mining in NET DOT framework	Setting and maintaining the accuracy of the hybrid model
[111]	2020	An IoT Honeynet Based on Multiport Honeypots for Capturing IoT Attacks	medium-high interaction	Capturing and analyzing attacks targeting IoT devices	N/A	Multi-port honeypots in a honeynet	Multi-port honeypot capability
[112]	2020	The Use of Honey-pot in Machine Learning Based on Malware Detection: A Review	Combined	Use of Honeybot in machine learning to detect malware	Techniques	Integration of honeypots with machine learning	effectiveness of detection systems.
[113]	2020	Enhanced attack blocking in IoT environments: Engaging honeypots and machine learning in SDN Open Flow switches	Combined	Improve attack blocking with ML techniques in Open Flow SDN switches.	Supervised	a Framework for integrating honeypots with Open Flow SDN switches and machine learning techniques.	Scalability

[114]	2021	Password Attack Analysis Over Honey-pot Using Machine Learning Password Attack Analysis	Low-interaction	Analyze password attacks.	Unsuper-vised	Use of honeypots and ML to analyze password attacks	Honey-pot maintenance
[115]	2022	Semi-supervised approach for detecting distributed denial of service in SD-honey-pot network environment	High interaction	Detecting distributed DDoS attacks	Semi-supervised.	Using honeypots and semi-supervised machine learning to detect DDoS attacks.	Data quality
[116]	2022	Predicting Attack Patterns via Machine Learning by Exploiting Stateful Firewall as Virtual Network Function in an SDN Network	High interaction	Predicting attack patterns in software-defined networks (SDN)	Supervised and unsuper-vised	Using honeypots and ML to predict attack patterns in SDN networks	Firewall with status as VNF

[117]	2022	Threat Prediction using Honeypot and Machine Learning	High-interaction	Predicting cyber threats	Supervised and unsupervised	Use of honeypots and ML for threat prediction and detection	Scalability and adaptability
[118]	2022	A Passive OS-Fingerprinting framework using Honeypot	Low interaction	Framework for identifying and classifying attackers' operating systems	Supervised and unsupervised	Use of honeypots and ML techniques for passive identification of SO	Active OS fingerprinting methods
[119]	2022	Honey Models: Machine Learning Honeypots	High-interaction	Use ML models to improve the effectiveness of honeypots in malicious detection and analysis	Clustering	Integration of ML techniques in high-interaction honeypots	Cost, adaptability and storage
[120]	2023	Analysis and Implementation of Honeypot Framework for Enhancing Network Security	High-interaction	Use of a honeypot framework that attracts, detects and analyzes malicious activity.	Supervised and unsupervised	Implementation of a high interaction honeypot framework	Storage and processing.

[121]	2023	Design of LAN Security Defense System Based on Honeypot Technology	High-Low-interaction	Design a security defense system for local area networks (LAN) using honeypot technology.	Supervised and unsupervised	Implementation of a honeypot defense system for LAN networks.	Storage and processing.
[122]	2024	Enhancing Network Security through a Multi-layered Honeypot Architecture with Integrated Network Monitoring Tools	High-Low-interaction	Improving network security through the use of a multilevel honeypot architecture	Supervised and unsupervised	Implementation of a multilevel honeypot architecture,	Advanced processing and storage capabilities.
[123]	2024	Enhanced honeypot security for intrusion detection and prevention systems using blockchain	High-interaction	Improving the security of intrusion detection and prevention systems (IDPS) using blockchain technology.	Supervised	Integration of blockchain with high-interaction honeypots	System complexity, costs and performance

Table 3.2: Types, objectives, contribution, and Limitations of Honeypots

The objectives of various research studies on honeypots have centered around detecting and analyzing attacks on systems. For example, Suleiman A. aims to identify network vulnerabilities and assess their resilience against attacks [118]. Similarly, [117], focuses on predicting vulnerable hosts within SDNFV networks with distributed controllers. Many studies also emphasize configuring and utilizing machine learning to predict attack mechanisms, threats, and intrusion detection [113], [114], [116], [117], [124], [125] On the other hand, [126] centers its research on medium-to-high interaction honeypots to analyze vulnerabilities like CVE-2017-17215, which is widely exploited by botnets [5], [24], [105], [107], [127]. In the information collected, several authors' research objectives were the design, architecture, and models that can be used to identify and profile attacks.

Regarding limitations, some studies were conducted in virtual environments. For instance [116], sought to predict susceptible hosts likely to be targeted in SDNFV networks with distributed controllers. This study highlighted a limitation: the accuracy of predictions decreased as the threshold for false positives increased. Furthermore, there is a notable absence of solutions applying Machine Learning techniques to the issues addressed.

The field of scientific research on honeypots in computer security has evolved significantly in recent years. High-interaction honeypots, now frequently deployed, offer a more realistic and effective simulation of network environments. Integration with existing security systems, such as firewalls, IDS, IPS, and WAF, has improved the interpretation of honeypot alerts. Although many studies still rely on statistical data analysis, there is a growing trend of applying Machine Learning techniques to enhance results and uncover patterns.

### **3.5. Cybercriminal Personality**

Research into the personality of hackers or cybercriminals remains relatively sparse. As seen in Figure 3.10, there was one publication in 2018, a slight increase of publications in 2020, and a similar number of publications in other years.

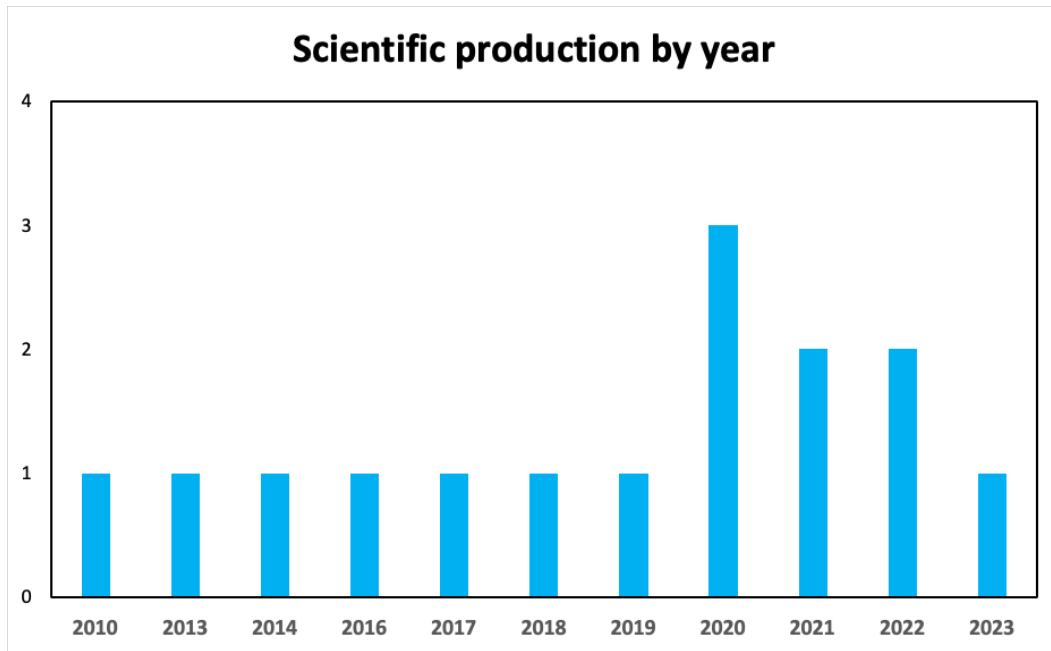


Figure 3.10: Primary studies, year of publications Cybercriminal Personality

In 2010, [128], conducted a quantitative study using a survey validated by experts. On the other hand, Summers et al. [129] employed qualitative research based on Grounded Theory, conducting semi-structured interviews with 18 hackers from a hacking community. Other studies have predominantly used quantitative descriptive and analytical methodologies, with some literature reviews included due to the limited number of studies in this area, as detailed in Table 3.3.

Ref	Year	Type	Study type	Title	Instrument	Scale	Population	Classification	Motivation
[130]	2010	Journal Article	Empirical	The Risk Propensity and Rationality of Computer Hackers	Surveys	Regression Models	Computer hackers	Established personality typologies and theories	Intrinsic and extrinsic factors
[129]	2013	Research Report	Qualitative	How Hackers Think: A Study of Cybersecurity Experts and Their Mental Models	Interviews and content analysis	Grounded Theory/ Triangulation	Cybersecurity experts	Mental models and psychological traits	Intellectual challenge, curiosity
[131]	2014	Book Section	Qualitative	The Psychology of Computer Criminals	Case studies	N/A	Convicted hackers and cyber-criminals.	Impulsivity and recognit	Gratification, revenge, profit, ideology and power and control
[132]	2016	Journal Article	Quantity	Hacker Personality Profiles Reviewed in Terms of the Big Five Personality Traits	Big five personality traits	Likert Scale	Six hacker subjects	-Hat white - hat black	Hacking activity

[133]	2017	Journal Article	Correlational	Computer criminal behavior is related to psychopathy and other antisocial behavior	Psychopathy Checklist-Revised (PCL-R)	Likert Scale	Hackers and cyber-criminals	Antagonism Emotional Stability Disinhibition Narcissism	Intellectual curiosity
[134]	2018	Journal Article	Correlational	Human resources and their tendency to information security crimes based on Holland's theory	John Holland's Theory	Statistical Analysis	Various organizations	According to Holland's theory	Power and control
[135]	2019	Conference Proceedings	Mixed	Youth hackers and adult hackers in South Korea: An application of cybercriminal profiling	Structured interviews	Likert scale	Young and adult hackers	Big Five Personality Traits	Curiosity, Recognition, Challenge, Revenge, Profit
[136]	2020	Conference Proceedings	Mixed	Psychological Profiling of Hacking Potential	Questionnaires and structured interviews	Likert scale	University students	Big Five Personality Traits	Curiosity, Recognition, Challenge, Revenge, Profit



[137]	2020	Conference Proceedings		Quantitative	Measuring Psychosocial Behavioral Factors Improves Attack Potential Estimates	Psy- and	Psychometric surveys	Likert scale	Attackers and security professionals	Big Five Personality Traits		Curiosity, Recognition, Challenge, Revenge, Profit
[138]	2020	Conference Proceedings		Quantitative	Predicting personality patterns of behavior collected with smartphones	per- from	Smartphone sensors	Likert scale	Smartphone users	Big Five		N/A
[139]	2021	Journal Article	Review		Profiling the Cybercriminal: A systematic review of research		Synthesis of previous studies	N/A	Studies reviewed	White hat Gray hat	Black hat	Economic, Reputational, Ideological, Curiosity
[140]	2021	Journal Article		Qualitative	Network discovery and scanning strategies and the Dark Triad	discov- and scanning	Questionnaires and structured interviews	Liker Scale	University students	Dark triad		narcissism and psychopathy

[141]	2022	Journal Article	Review	Are you anonymous? Social-psychological processes of hacking groups	N/A	psychological research	N/A	-criminals warriors -hacktivists -coders	-cyber -hack- -insiders -coders	Ideology, prestige, recreation, and revenge
[142]	2022	Conference Proceedings	Review	The Amorphous Nature of Hackers: An Exploratory Study	Hacker Perception Questionnaire	Liker Scale	University students	-White -Black -Gray hacker	hacker hacker	Ideology, prestige, recreation, and revenge
[143]	2023	Journal Article	Mixed	Comparing cyber offenders and offline offenders on HEXACO personality domains and their underlying facets	Questionnaire	Liker scale	individuals	Cyber offenders		Ideology, prestige, recreation, and revenge

Table 3.3: Methods, Instruments, and Personality Classification of Hackers or Cybercriminals

The primary instrument utilized in various studies on cybercriminal personality is the FBI criminal profiling framework [135]. On the other hand, most studies use the Big Five Personality Traits test composed of 132 items, is commonly employed [132], [135], [138]. This test evaluates key personality dimensions such as openness, conscientiousness, extraversion, agreeableness, and neuroticism. Al-Ajilouni [134] used John Holland's Career Choice Theory (RIASEC), a 66-item instrument traditionally used for career assessment but also effective in evaluating personality traits. Similarly, Seigfried-Spellar et al. [133], applied the Elementary Psychopathy Assessment (EPA), a 178-item self-report measure designed to assess fundamental elements of psychopathy.

Bachmann [129] conducted personality assessments at the Shammon-Con hackers' annual convention, targeting network professionals and hackers selected based on specific criteria. Matulesy et al. [132] used six hackers for their research. Seigfried-Spellar et al. [133] conducted their study on 250 Internet users who may or may not be considered hackers. In [129], the personality traits of 83 cybercriminals in a South Korean prison are evaluated. Consequently, in the research [138], [143], [144], they measure the potential profile of cyber attackers in students of some technical careers and individuals in general.

The various instruments used to assess hacker or cybercriminal personalities have led to several classification schemes based on personality traits. Notably, the repeated classifications include White Hat, Black Hat, and Gray Hat hackers [132], [139], [142], [144]. Motivations for cybercriminal behavior include revenge, boredom, ideology, ego, sabotage, espionage, gratification, and blackmail. These motivations are illustrated in Figure 3.11 which provides a diagram of the different motivational factors driving cybercriminal activities.



Figure 3.11: Word diagram of the motivation of cybercriminals

A review of hacker personality research reveals its interdisciplinary nature, encompass-

ing psychology, computer security, and criminology. However, it is essential to acknowledge that a singular or stereotypical profile cannot encapsulate all hackers, as they exhibit diverse motivations and characteristics. The personality of a hacker can evolve and be influenced by various contextual and social factors.

Hackers typically possess high technical proficiency, including skills in programming, computer networks, and operating systems. Their curiosity and intellectual drive often lead them to self-teach and challenge existing technical barriers. Hackers exhibit creative and divergent thinking, enabling them to devise innovative solutions to complex problems.

Motivations for hacking vary widely and include financial gain, recognition, curiosity, activism, or the desire to cause disruption. Hackers often show a low tolerance for authority and established norms, with some experiencing heightened emotions like excitement or adrenaline during attacks. It has proven challenging to categorize hackers into a single profile or sample, leading researchers to explore diverse data sources, including forums, hacker communities, and technology professionals. Psychological instruments used to study hacker behavior include those designed to identify behavioral patterns, with the Big Five Personality Traits model being particularly effective in classifying hackers into categories such as White Hat, Gray Hat, and Black Hat.

The Big Five Personality Traits model, a widely accepted psychological framework, describes five core dimensions of human personality:

1. **Openness to Experience:** This trait reflects a person's willingness to explore new ideas, experiences, and emotions. Individuals high in openness are imaginative, curious, and adaptable.
2. **Conscientiousness:** This dimension pertains to an individual's organization, responsibility, and reliability. Those high in conscientiousness are disciplined, orderly, and follow through on commitments.
3. **Extraversion:** This trait measures a person's desire for stimulation and social interaction. Extraverts are typically energetic, assertive, and enjoy being around others.
4. **Agreeableness:** This dimension indicates a person's tendency to be empathetic, cooperative, and considerate toward others. Agreeable individuals have a positive and compassionate disposition.

5. Neuroticism: This trait involves the degree to which a person experiences negative emotions such as anxiety, emotional instability, or stress.

Table 3.4 provides a detailed overview of hacking definitions, motivations, messages left behind, and justifications. This table offers insights into:

Definitions: Key terms and concepts related to hacking and cybersecurity.

Motivations: The underlying reasons driving hackers, whether financial, political, social, or personal.

Messages Left Behind: The types of messages or signals hackers leave post-attack, including claims, threats, or warnings.

Justifications: The rationales hackers provide for their actions, ranging from protesting injustices to demonstrating technical skill.

This detailed analysis helps to contextualize the driving factors behind hacking activities and the methods employed by hackers.

<b>Ref</b>	<b>Motivation</b>	<b>Definition</b>	<b>Message</b>	<b>Justify</b>
Oxford Dictionary [131], [143]	Hackivism Or Political	Hackivism involves malicious acts on the Internet to promote political, religious, or social ideas, using electronic devices to defend their ideals.	Loved By Linda Long Life Vietnam   Moroccan Revolution	The hacker leaves a message about a social, political, or religious problem.
Oxford Dictionary [131], [133]	Ego	A person's sense of self-esteem or self-importance.	LapanWasTaken Whoops.Got Hacked	Here They leave messages to showcase skills and talents that set them apart.
Oxford Dictionary [131], [135], [142], [143]	Revenge	Action of causing harm to someone in retaliation for an offense or grievance received.	hacked by Salim Alk, ohh, sorry your security is gay	The hacker leaves a message with their identifier, typically mocking the security of the hacked site.
Oxford Dictionary [133], [143]	Entertainment	The action of providing or being provided with amusement or enjoyment.	Hacked by Phenix-TN Just for fun, HAHAAHHA! ANYTIME I LIKE TO LOL, THANKS TO IMAM	A hacker who performs cyberattacks just for fun
[131], [136], [130]	Monetary	Connected with money	Hacked By Babacang07 - PhantomSec1337, icq: Gh05t11n6, telegram: Flavvy7	The only interest is for monetary gain, and he leaves his data to be contacted for data recovery or security patches.

Oxford Dictionary [131], [136], [137]	boredom	Feel weary because one is unoccupied or lacks interest in one's current activity.	Hacked By Ahd, Hacked By Ahd, This world is bad	A hacker who performs cyber-attacks to pass the time and be engaged in some activity other than boredom.
Oxford Dictionary /Google [145], [131], [129], [136], [141]	Recognition	To be recognized for an act or action that has been performed involves a great sacrifice of either intelligence or time.	Hacked by Mr.kro0oz.305	He wants to be recognized as the one who attacked the system, leaving a basic message of his Nickname.

Table 3.4: Big Five Personality Traits

This table provides a comprehensive view of the dynamics behind hacking, facilitating a better understanding of this complex phenomenon.



## 4. PATTERN CLASSIFICATION

### Contents

---

4.1	Materials and methods . . . . .	47
4.2	Web-Attacks Cognitive Patterns Classifier Architecture . . . . .	49
4.3	Analysis of the infrastructure T-Pot Honeypot . . . . .	51
4.4	Analysis of studies on the personality of hackers . . . . .	59
4.5	Application of the BFQ Survey . . . . .	63

---

### 4.1. Materials and methods

This section outlines the methodology employed in this dissertation, detailing the materials and tools utilized to develop the research proposal. The methodology integrates both experimental and non-experimental research approaches, specifically employing honeypot systems and utilizing a descriptive, correlational, cross-sectional cohort design. This comprehensive approach enables a thorough examination of the personality traits and behaviors of potential cyber attackers. The research is structured into the following phases:

#### Phase 1 Information Gathering

- Review existing literature on cyber attackers' personality traits and motivations.
- Examine the effectiveness of honeypot systems in catching cyber attackers.  
Examine survey methods used in cybersecurity studies.

#### Hypothesis Formulation

- Formulate research hypotheses based on the literature review.
- Identify key personality traits that may be associated with cyber attackers.
- Develop a conceptual framework linking personality traits to cyberattack behaviors.

## **Phase 2 Experimentation**

Honeypot system design and implementation

- Design a honeypot system that simulates vulnerable networks or systems to attract potential cyber attackers.
- Select appropriate honeypot software and hardware based on the study's objectives.
- Deploy honeypot systems strategically to capture a diverse sample of cyber attackers.

Collect data from the honeypot system:

- Monitor and record the attackers' activities interacting with the honeypot system.
- Collect data on attackers' methods, techniques, and patterns.
- Store collected data for further analysis.

## **Phase 3 Descriptive Design** Survey design

- Determine instrument to assess personality traits of potential cyber attackers through literature review.

Survey Distribution and Data Collection

- Distribute the survey to potential respondents, ensuring anonymity of participants, using techniques such as social media postings and online forums.
- Collect survey responses and securely store data for analysis.

## **Phase 4 Analysis of Results**

Data Analysis:

- Analyze honeypot data using descriptive and inferential statistics to identify common patterns and trends (machine learning).
- Analyze survey data to assess the relationship between personality traits and cyberattack behavior.
- Perform correlation and regression analysis (machine learning classification) to explore the strength and direction of these relationships.
- Compare the results of the honeypot and survey data to validate the hypotheses.

**Results and Discussion:**

- Summarize the main findings of the study and discuss
- their implications for understanding personality hackers. Compare the results to existing literature and explain any discrepancies.
- Discuss the study's limitations and suggest areas for future research.

**Conclusion:**

- Summarizes the study's objectives, methods, and findings.
- Provide recommendations for using the results to improve cybersecurity efforts and deter potential cyber attackers.
- Emphasize the importance of continued research to understand the personality traits of cyber attackers for more effective prevention and response strategies.

**4.2. Web-Attacks Cognitive Patterns Classifier Architecture**

This research presents a theoretical architecture designed to protect web servers using data obtained from honeypots [146]. As a first step, a T-POT with WCPC will be deployed in hacker communities to collect information about different attack techniques. The honeypots will collect data that will be used to analyze the cognitive patterns of attackers. This data will be analyzed using machine learning techniques to identify the most effective pattern classification method. The machine learning algorithm can be either clustering or unsupervised principal component analysis (PCA). Several characteristics of the attacks are considered when analyzing the data, such as time zone (although attacks usually use intermediate machines such as VPN and botnets), attack technique, fingerprint removal techniques, and keystrokes. As shown in Figure 4.1, the WCPC consists of the following elements:

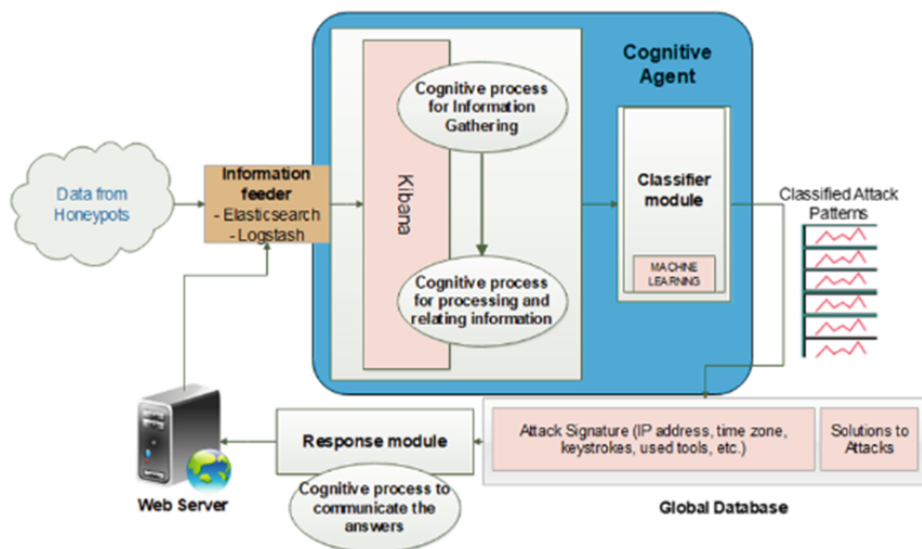


Figure 4.1: WPCP architecture

**Information Feeder:** This component is responsible for retrieving data from the Internet using Elasticsearch, allowing the creation of a multi-trend search engine. It also uses Logstash to classify the information from the logs coming from the honeypots.

**Cognitive Agent:** This module includes cognitive processes aimed at collecting, processing and correlating information, as well as communicating the results. It uses machine learning techniques supported by Kibana to analyse the data collected by the information feed. The processed information will contain the cognitive patterns of the attackers. Once these patterns have been evaluated, the information is stored in the global database.

**Global Database:** All attack information, once classified, will be stored in this database according to its signatures, such as the attacker’s IP address, time zone, keystrokes, tools used and traces left behind. In addition, the database will contain other relevant data such as the URLs of the hacked sites (extracted from sites that publish this information, such as <http://www.zone-h.org>) and the solutions/responses to the attacks.

**Response Module:** This module implements actions to prevent or mitigate attacks, using previously known solutions for each type of attack. Thanks to the classification of cognitive patterns, the type of attack can be identified. Each response is recorded in the global database to document system activity.

The above elements interact through the following processes:

**Pre-classification:** This process, performed by the information feeder, is responsible for collecting attack data from honeypots and the Internet. Two open-source tools are used in this process: (1) Logstash, which collects device logs for future searches, and (2) Elastic-

search, which contributes to the cognitive process by searching data that is being modified in real time.

**Cognitive process:** This process analyses the collected information using a number of parameters. Kibana, an Elasticsearch data exploration, visualization and discovery tool, supports this processes by classifying information according to specific parameters. For example, if the attacker performs a port read (using tools such as Nmap, nbtscan and Metasploit) during the fingerprinting phase, this activity is recorded in the logs, and the classifier module associates the extracted information with a given classification pattern.

**Classification process:** This process is carried out by the classifier module, which is responsible for the final evaluation of the information. The evaluated information is stored in the appropriate database according to its classification.

**Response Process:** This process is activated once the Global Database is operational. When a WCPC web server receives an attack that is already known to the system, the response module executes the solutions stored in the global database.

The ELK stack toolset (Elasticsearch, Logstash and Kibana) was chosen for its ability to ingest and analyses data in real time. This toolkit is very useful for the project as it allows the collection and analysis of honeypot logs. Another advantage of using these tools is that they are open source, which allows them to be modified with new features or adaptations as needed.

### **4.3. Analysis of the infrastructure T-Pot Honeypot**

For the execution of this project, an infrastructure was implemented dedicated to the purpose that consisted of a first firewall owned by the service provider CEDIA that, together with the ESPOCH, managed the creation of a tunnel to protect the data infrastructure and services of the Higher Education Institution. The server used for the purpose was an HP ProLiant DL360 Gen9; the same was located in the Faculty of Informatics and Electronics FIE in the ESPOCH, and the installation process of the T-Pot was developed, which previously was downloaded. ISO image from GitHub for subsequent installation. Afterward, all the services that ESPOCH, such as its website and databases, were replicated. The use of the different open ports in Table 4.1 on a web server was also guaranteed, in such a way, incentivizing the attacker to perform different attacks such as DDoS, XSS, SQL injection, brute force, CSRF, and buffer overflow attacks.

Service	Port	Description
HTTP	80	Hypertext Transfer Protocol
HTTPS	443	Secure HTTP version
FTP	21	File Transfer Protocol
SSH	22	Remote Shell
SMTP	25	E-mail protocol (Send)
POP3	110	E-mail protocol (Receive)
IMAP	143	E-mail Management
MySQL	3306	Database Management
PostgreSQL	5432	Database Management
MongoDB	27017	Database Management
HTTP alternativo	8080	Alternate port HTTP web traffic

Table 4.1: Description of protocols

The server addresses were published on forums and cyber-criminal or hacker sites. The system has been online for more than six months. The information is collected from honeypots that have the most interactions with the network, such as Dionaea, Nginx, Adbhoey, and Ciscoasa. The collected data is then analyzed to present its results.

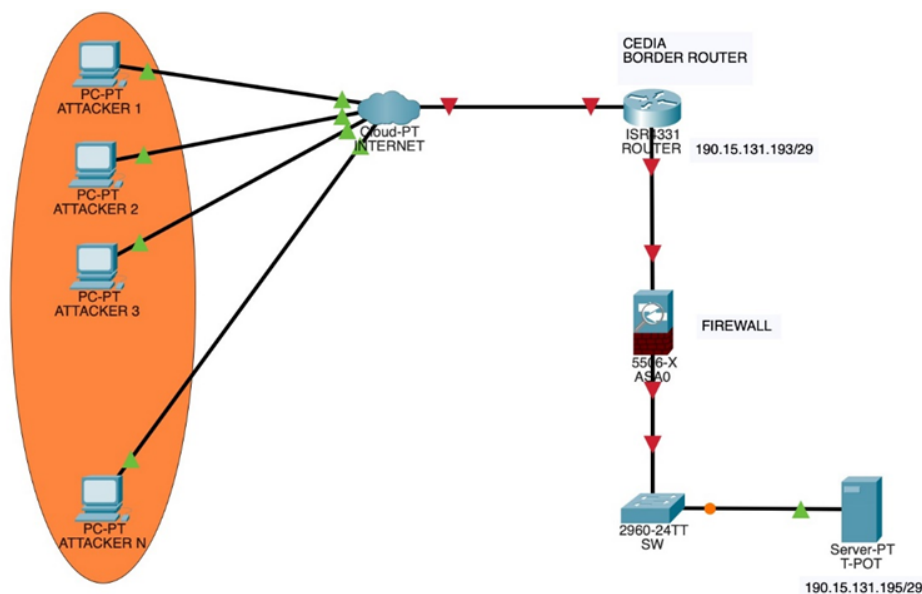


Figure 4.2: Server infrastructure

For the data collection, as shown in Figure 4.2, the infrastructure of the Escuela Politécnica del Chimborazo was used, and as mentioned above, the bandwidth was provided by CEDIA. In the first instance, being inside an educational institution, the first attacks reached other devices and servers, so creating a direct channel with the server was necessary, thus avoiding exposing the institution's own devices, data, and servers.

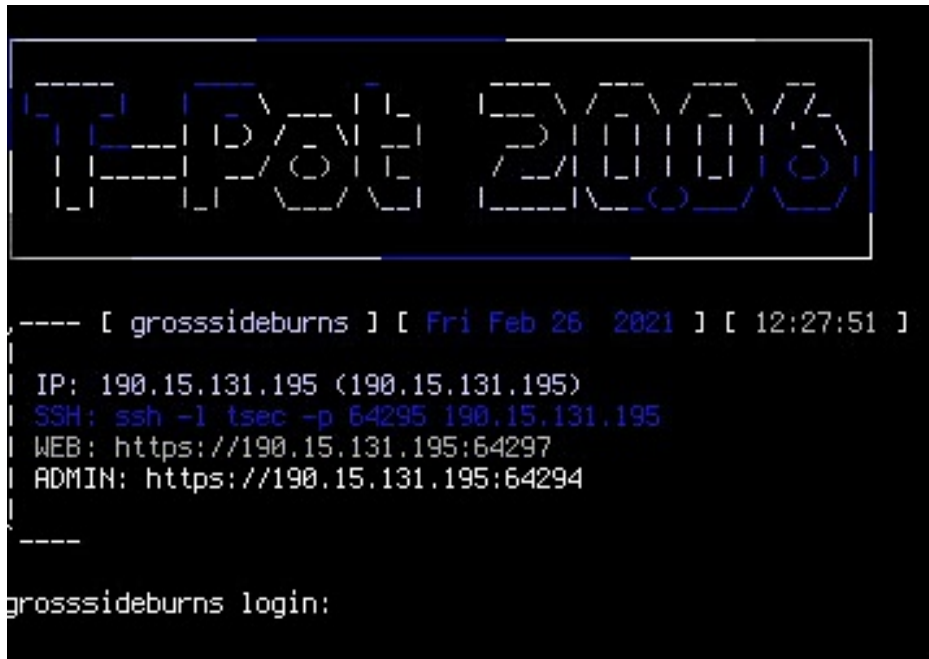


Figure 4.3: Login T-Pot Honeypot

The T-Pot Honeypot was administered through the server address from an SSH and HTTPS connection. Figure 4.3 shows the server's main screen; through the ELK Stack, graphical administration and visualization of the results become super simple. However, a high level of knowledge is necessary for additional or personal configurations according to the purpose of the research. On the other hand, Figure 4.4 shows the server activity with the active ports of the different services mounted on the Honey Pots hive.

```

190.15.131.195 Terminal
[root@grossideburns:/opt/tpot/bin]# ./dps.sh
===== System |=====
Date: Sun 28 Feb 2021 10:29:08 PM UTC
Uptime: 22:29:08 up 1 day, 10:01, 0 users, load average: 2.97, 2.48, 2.34

Registros
NAME STATUS PORTS
Almacenamiento adbhoney Up 22 hours 0.0.0.0:5555->5555/tcp
ciscoasa Up 22 hours
citrixhoneypot Up 22 hours 0.0.0.0:443->443/tcp
Redes conpot_guardian_ast Up 22 hours 0.0.0.0:10001->10001/tcp
conpot_iec104 Up 22 hours 0.0.0.0:161->161/tcp, 0.0.0.0:2404->2404/tcp
conpot_ipmi Up 22 hours 0.0.0.0:623->623/tcp
Contenedores conpot_kamstrup_382 Up 22 hours 0.0.0.0:1025->1025/tcp, 0.0.0.0:50100->50100/tcp
Cuentas cowrie Up 22 hours 0.0.0.0:22-23->22-23/tcp
cyberchef Up 22 hours (healthy) 127.0.0.1:64299->9000/tcp
dicompot Up 22 hours 0.0.0.0:11112->11112/tcp
Servicios dionaea Up 17 hours 0.0.0.0:20-21->20-21/tcp, 0.0.0.0:42->42/tcp,
0.0.0.0:81->81/tcp, 0.0.0.0:135->135/tcp, 0.0.0.0:445->445/tcp, 0.0.0.0:1433->1433/tcp, 0.0.0.0:1723->1723/tcp, 0.0.0.0:1883->1883/tcp, 0.0.0.0:3306->3306/tcp, 0.0.0.0:69->69/udp, 0.0.0.0:5060-5061->5060-5061/tcp, 0.0.0.0:27017->27017/tcp, 0.0.0.0:5060->5060/udp
Actualizaciones de elasticpot Up 22 hours 0.0.0.0:9200->9200/tcp
Software elasticsearch Up 22 hours (unhealthy) 127.0.0.1:64298->9200/tcp
ewsposter Up 16 hours
Aplicaciones fatt Up 22 hours
head Up 22 hours (healthy) 127.0.0.1:64302->9100/tcp
heralding Up 16 hours 0.0.0.0:110->110/tcp, 0.0.0.0:143->143/tcp, 0.0.0.0:993->993/tcp, 0.0.0.0:995->995/tcp, 0.0.0.0:1080->1080/tcp, 0.0.0.0:5432->5432/tcp, 0.0.0.0:5900->5900/tcp
honeysap Up 22 hours 0.0.0.0:3299->3299/tcp
honeytrap Up 22 hours
kibana Up 22 hours (healthy) 127.0.0.1:64296->5601/tcp
logstash Up 22 hours (healthy)
mailoney Up 22 hours 0.0.0.0:25->25/tcp
medpot Up 22 hours 0.0.0.0:2575->2575/tcp
nginx Up 22 hours
p0f Up 22 hours
rdpy Up 22 hours 0.0.0.0:3389->3389/tcp
snare Up 22 hours 0.0.0.0:80->80/tcp
spiderfoot Up 22 hours (healthy) 127.0.0.1:64303->8080/tcp
suricata Up 22 hours
tanner Up 22 hours
tanner_api Up 22 hours
tanner_phpox Up 22 hours
tanner_redis Up 22 hours 6379/tcp

```

Figure 4.4: Honeypot initialized ports

Among the services mounted on the web server, the main web page of the ESPOCH was replicated. The Httrack tool was used; first, the information was downloaded from the website, which took approximately 36 hours. Afterward, the web server was previously mounted on the physical server implemented, as shown in Figure 4.5



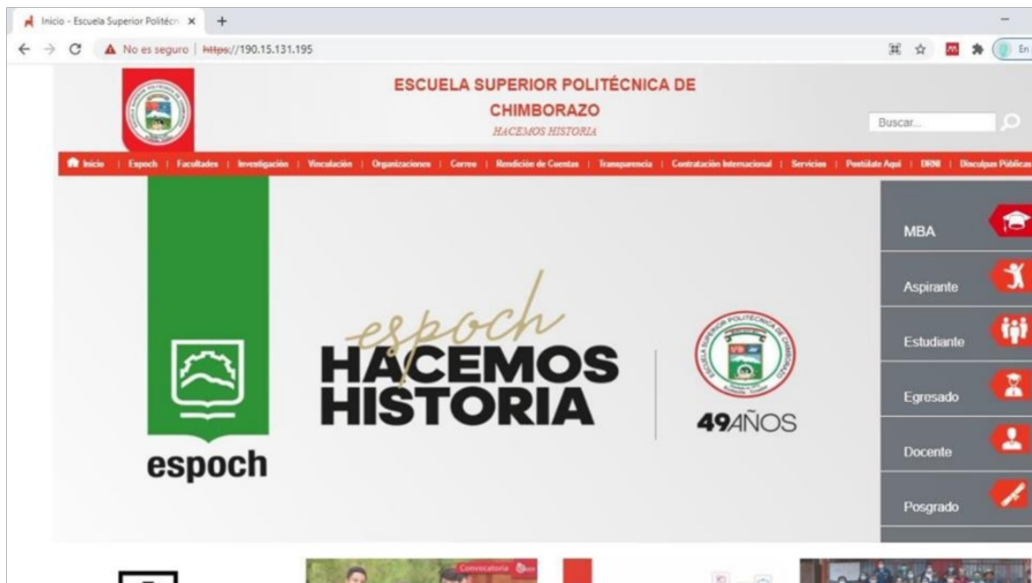


Figure 4.5: IP of the ESPOCH homepage

Among the main results after more or less six months of the server being implemented in the ESPOCH infrastructure, it was observed that the attacks perpetrated by DoS occupy the first place, with a large number of attacks of 39%. They were followed by the SQL injection attack at 21% and cross-site scripting (XSS) at 12%. Similarly, with 10% Brute Force Attacks, cyber-criminals prefer attacks as observed in a previously conducted study [147], shown in Figure 4.6.

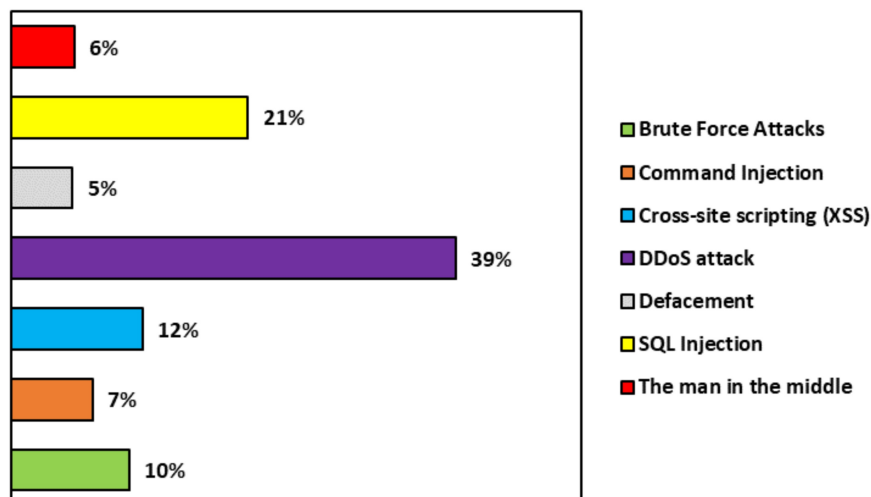


Figure 4.6: Attacks on the server implemented in the ESPOCH infrastructure

Also, an important data that can lead to valuable conclusions from this study is the time

zone in Figure 4.7. The T-Pot Honeypot recorded 18 time zones as favorites, among which in the first place is the GMT-7 zone, with 26.6% belonging to Asian countries such as Indonesia, Thailand, Cambodia, and Singapore. Then comes GMT+7, representing the USA with 24.8%. They were followed by 12.4% for GMT+2 European countries such as Germany, France, Serbia, Netherlands, and India. In the same way, with 7% for GMT+3 Turkey, Israel, Iran Islamiza, Syria, and GMT-3 Brazil. It should be considered that most cyber-criminals use VPN; however, it can also be a pattern that allows them to improve the security of a web server since the time zones are repeated, which will enable them to determine that they do not change VPN constantly, but rather hide the address of origin, but are not permanently randomizing the IPs used for attacks.

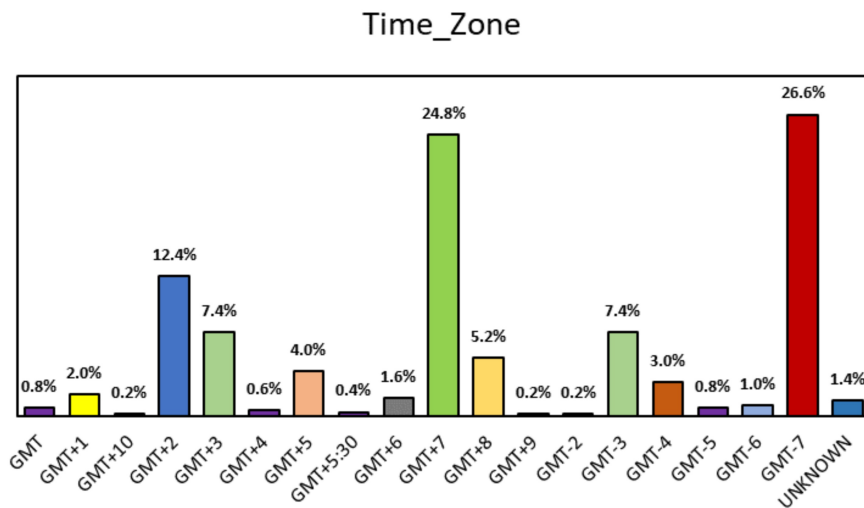


Figure 4.7: The time zone from which the Attacks were launched

The T-Pot allows configuring a farm of servers in addition to the web above page. The results in Figure 4.8 show a preference for attacking web servers with Apache 46.20% of the attacks against this server. Among the data recorded in the T-pot, Honeypot is 32.20% to more than two servers recorded with the unknown label. Next, with 12.60% nginx, the latter server has become popular in the last decade due to its robustness.

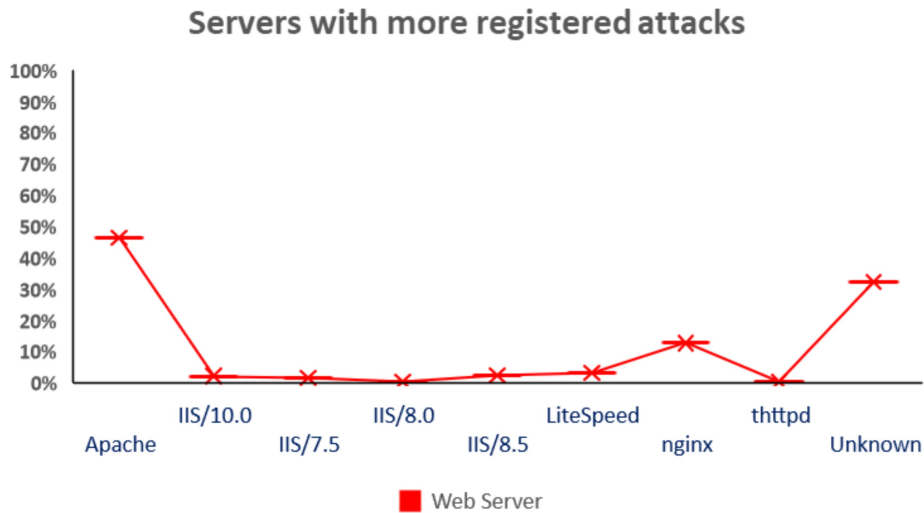


Figure 4.8: Server with more registered attacks

As previously discussed, cyberattacks are not confined by nationality or geographic location, making it challenging to pinpoint the exact origin of these attacks. The use of IP masking tools and VPNs further complicates tracking the true source of cyber threats. Despite these challenges, certain countries, such as the United States, China, Russia, North Korea, and Iran, are frequently implicated in significant cyber espionage and large-scale attacks. Additionally, some developing nations are also suspected of engaging in illicit cyber activities. According to Figure 4.9 data from the T-Pot system indicates that the United States experiences the highest number of attacks, followed by Indonesia, Brazil, and Iran.

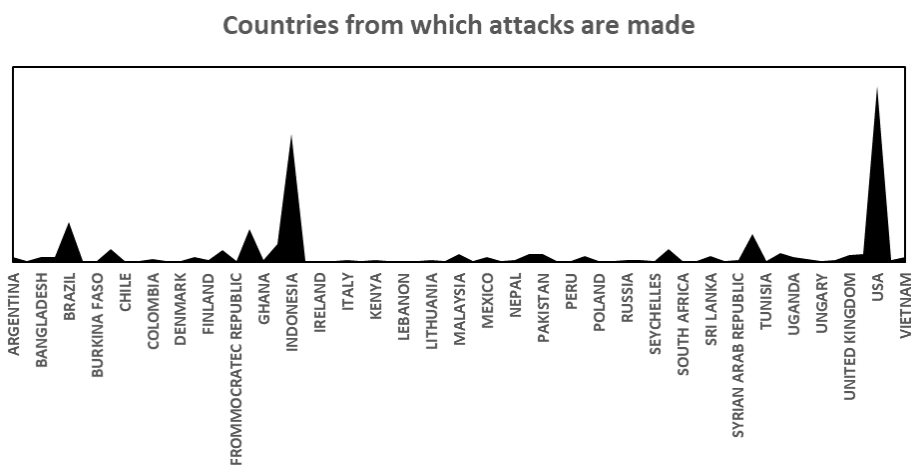


Figure 4.9: Countries from which attacks are made

In the research of Back et al. [110], it is stated that ego is fourth on the list of cyber-criminals' motivations, especially when they are part of a group or seek to demonstrate

their leadership. The T-Pot honeypot allows the digital fingerprints of cybercriminals to be recorded through their signatures or nicknames (Nicknames), and most of them leave their rubric. This phenomenon is illustrated in Figure 4.10, which presents crucial data for detailed profiling of cybercriminals.

In addition, cybercriminals expose these nicknames on numerous forums to develop an on-line reputation. In the realm of web security, this data can be extremely useful in associating criminals' activities with their profiles, thus enabling the implementation of more effective security measures based on this information. More sophisticated and proactive defense strategies can be developed by analyzing and correlating the data collected by honeypots such as T-Pot with the information available in these forums.

Mr.Kro0oz.305	Mr.Kro0oz.305	Tahun FROMpan Nikal Ayyıldız Tim	./Anon666Txploit	P4r4d0x Cr3w	Mr.Z
EbRaHiM-VaKeR	Mr.Kro0oz.305	LahBodoAmat	Mr.Ti74N	Mr.Och4n	Black_TeamX
SABUNMANDI CYBER TI	Mr.Kro0oz.305	finistro lammer	Crystal_MSf	AKINCILAR	evil-net
Admereal zino_dz	Mr.Kro0oz.305	YIIX103	NULL SQU4D	KAKEGURAI	T-Freak
Hamza Anonime	Faisal 1337	aDriv4	Salim Alk	pi.hack	IDOLSEC Team
Black_X12	theMx0nday	finistro lammer	Akbar dravinky	Trenggalek Cyber Army	Unknown AI
Sofian X35 Dz	Mr.ToKeiChun69	Imam	DR1D1	r1dA	Bloos3rpent
LahBodoAmat	White Cyber Illusion	MiSh	PohonSagu	R0cket	Mister_XID
FoursFROMMath Team	Moroccan Revolution	TUNOVATO	Royal Battler BD	Yemen Cyber Army	chinfans
AlfabetoVirtual	Ox1998	Clash Hackers	R13S	Simsimi	Zer0FaulT
NDA	KrdSec	iccamy	AnoaGhost	LucasOwna	djebbaranon
Collapse Gang	Family Attack Cyber	NDA	Tn Jones	RexShelby	Mr.Jenskins
pr0s3x	Wedus	Ghost Hunter Illusion	RzkyO	Banjarnegara Xploit	dhuua
Fatal Error	CLAN_X12	TurkHackTeam	AnoaGhost	Hadii666h05T	Zodi4cxs
m1kesecurity	ro0t-M8n	Umam1337	LahBodoAmat	Zyyy	Mr.XaaD
Black_X12	Paraná Cyber Mafia	D3D0T	Zeerx7	B3g0k[Kurdish Hacker]	FRK48
KrdSec	D.R.S Dz Team	sh007	K4TSUY4-GH05T	V1ruz3L	VenoRyan
Arch1999	Panataran	./MrTahuSumedang	M15T4k3	Typical Idiot Security	Black-Python
Mr.Kro0oz.305	/Rayzky_	Toro	Trenggalek Cyber Army	F.Z Malai katHati	ALHOSANE
mr.anFROMrson	Sc0rp10n.DZ	AnonSec Team	Salman Hacker	./s3nt1n3L	FurkHan
K4PUYU4K	Ren4Sploit	z3ran gaza hack3er	./Cyber00t	Mamad Warning	Panataran
Stoupid!	Fuck MALAYSIA	Indonesia Attacker #0	Jebe Was Here!	Hii	SH3LL_INS3RT3R
Mosawi--	bky992	CyberFrost	Matigan1337	./Newbie4rt_ID	./Juba_Dz
Babacang07	oTm4n3	Mr.Froggy	./KeyzNet	RootAyyıldız	Centra7x
Melody-x48	Clash Hackers	luxe	KosameAmegai	./Synchronizer	FroxssHunter
Ahd404	SDM21	LucasOwna	Al Catraz	Illusion Silent Killer	bilat
TheLevelSevenCrew	Cubjmet7	TheZero	ZERA	AhmadBLocker	1975 Team
KURD ELECTRONIC TEA	Lumajang Xploit	z7F HaCkEr	Frustrated Hackers B	Oxloshua	Albania Attacker
./unn0maL	Unknown AI FUCKTARE	Galang10	Tn Bekicot	Astra	Verad
7870a	Shield iran	Dhen Bhocil	VaporEKJet	./Mr.xWanz403x	Hewmen
					Mr#

Figure 4.10: Database of Nicknames of Cybercriminals

In the project's first phase, a T-Pot honeypot server was implemented in the Escuela Superior Politécnica del Chimborazo (ESPOCH) infrastructure. This server was installed to attract and record malicious activity for later analysis. The Corporación Ecuatoriana para el Desarrollo de la Investigación y la Academia (CEDIA) provided the necessary bandwidth to ensure optimal honeypot operation. This collaboration made it possible to establish a solid foundation for the study and monitoring of cyberattacks, facilitating the collection of valuable data that will contribute to improving the institution's cyber defense strategies. In addition, the implementation of this honeypot in an educational environment provides a practical re-

source for the training of future professionals in the field of cybersecurity. [148].

Subsequently, a survey was developed based on the personality traits of the Big Five model, consisting of 132 personality-related questions and seven specific questions about the data obtained by the honeypot. This survey was shared on the same sites where the IP address of the T-Pot honeypot was located, as mentioned above.

The data collected was consolidated into a database structured in 18 columns by 500 rows. This database included detailed information such as the type of attack, the tool used, the operating system, the IP address, the hacker's personality, his nickname, and any messages left, among other characteristics.

Using a hybrid honeypot, i.e., a low-interaction and high-interaction honeypot, made it possible to simulate near-real computing environments. T-Pot, an open-source honeypot system, can be adapted to the needs and requirements of different network infrastructures. This project sets an important precedent by demonstrating that tools such as T-Pot can be used for scientific research to understand better the tactics, techniques, and tools that cybercriminals employ. In addition, this information is crucial for strengthening the security of computer networks.

The implementation and analysis of this honeypot provided valuable data on cyberattacks and a solid foundation for future cybersecurity research, contributing to developing more robust and adaptive strategies to protect critical infrastructures.

#### **4.4. Analysis of studies on the personality of hackers**

Based on a review of hacker personality studies, detailed in Table 4.2 are the methods, tools, and classifications used to assess the personality of hackers or cybercriminals. Tools employed for this purpose include the FBI Criminal Profiling Framework, John Holland's Career Choice Theory (RIASEC), the Elementary Psychopathy Assessment (EPA), and the Big Five Questionnaire (BFQ), among others.

This analysis reveals that the BFQ is the most accurate tool for measuring cybercriminals' personality and behavioral patterns. The Big Five Questionnaire (BFQ) comprehensively assesses personality traits along five key dimensions: openness to experience, responsibility, extraversion, agreeableness, and neuroticism. This accuracy allows for a more detailed understanding of the psychological factors that motivate cybercriminals and facilitates more accurate profiling, as shown in Figure 4.11.

In addition, combining these tools with empirical data obtained from honeypots, such as the

T-Pot, allows for a deeper correlation between observed behaviors and personality traits. This not only improves the ability to predict and prevent future attacks but also contributes to developing more effective intervention strategies tailored to attackers' specific psychological profiles.

<b>Instrument</b>	<b>Approach</b>	<b>Dimensions evaluated</b>	<b>Evaluation structures</b>	<b>Theoretical basis</b>	<b>Use</b>
BFQ	Big Five	Openness, responsibility, extraversion, friendliness, emotional stability	Continuo	Big Five Theory	General personality assessment in various contexts
MMPI	Psychological disorders	Various dimensions associated with mental health	Predefined categories	Clinical and psychology	Clinical and health psychology
MBTI	Personality types	Extraversion/introversion, perception/judgment, others	Predefined categories	Jungian Theory	Three-factor theory
EPQ	Three main dimensions	Extraversion/introversion, neuroticism, psychoticism	Scoring on specific scales	Personal development, team development	Psychological and clinical research

Table 4.2: Comparison of Big Five with other instruments

The BFQ has been used in psychological research and clinical practice to understand better human personality and its impact on various aspects of life. Its approach and five dimensions make it a valuable tool for exploring individual differences and understanding how they influence people's behaviors and experiences in different contexts.

The BFQ (Big Five Questionnaire) is designed to assess and understand an individual's fundamental psychological traits by measuring five core dimensions of personality: openness to experience, conscientiousness, extraversion, agreeableness, and emotional stability. Unlike other personality assessments, the BFQ uses a Big Five framework to provide a nuanced profile of personality traits along a continuum rather than categorical types or specific behavioral patterns.

In contrast, the Minnesota Multiphasic Personality Inventory (MMPI) is geared towards identifying psychological disorders and mental health issues. It focuses on detecting behavioral patterns linked to psychopathology rather than evaluating distinct personality traits such as extraversion or agreeableness.

The Myers-Briggs Type Indicator (MBTI), derived from Jungian theory, categorizes individuals into one of 16 personality types based on dichotomous preferences, such as Extraversion/Introversion and Perception/Judgment. Unlike the BFQ, which evaluates personality traits on a continuum, the MBTI assigns individuals to specific categories.

The Eysenck Personality Questionnaire (EPQ) assesses three primary dimensions: extraversion, neuroticism, and psychoticism. While it shares similarities with the BFQ in measuring basic personality traits, the EPQ employs a different scoring structure and focuses on a more limited set of dimensions.

Each of these instruments—BFQ, MMPI, MBTI, and EPQ—contributes uniquely to the field of psychology through its theoretical framework, measured dimensions, and scoring methods. This diversity in approaches ensures that these tools are tailored to different research objectives and applications.

For instance, the BFQ captures personality traits on a spectrum, such as openness to experience ranging from highly imaginative and open to new ideas to more conventional and routine-oriented. Extraversion in the BFQ spans from highly sociable and adventurous to more reserved and introverted. Table 4.2 provides a comparative overview of these instruments, highlighting their unique features and applications.

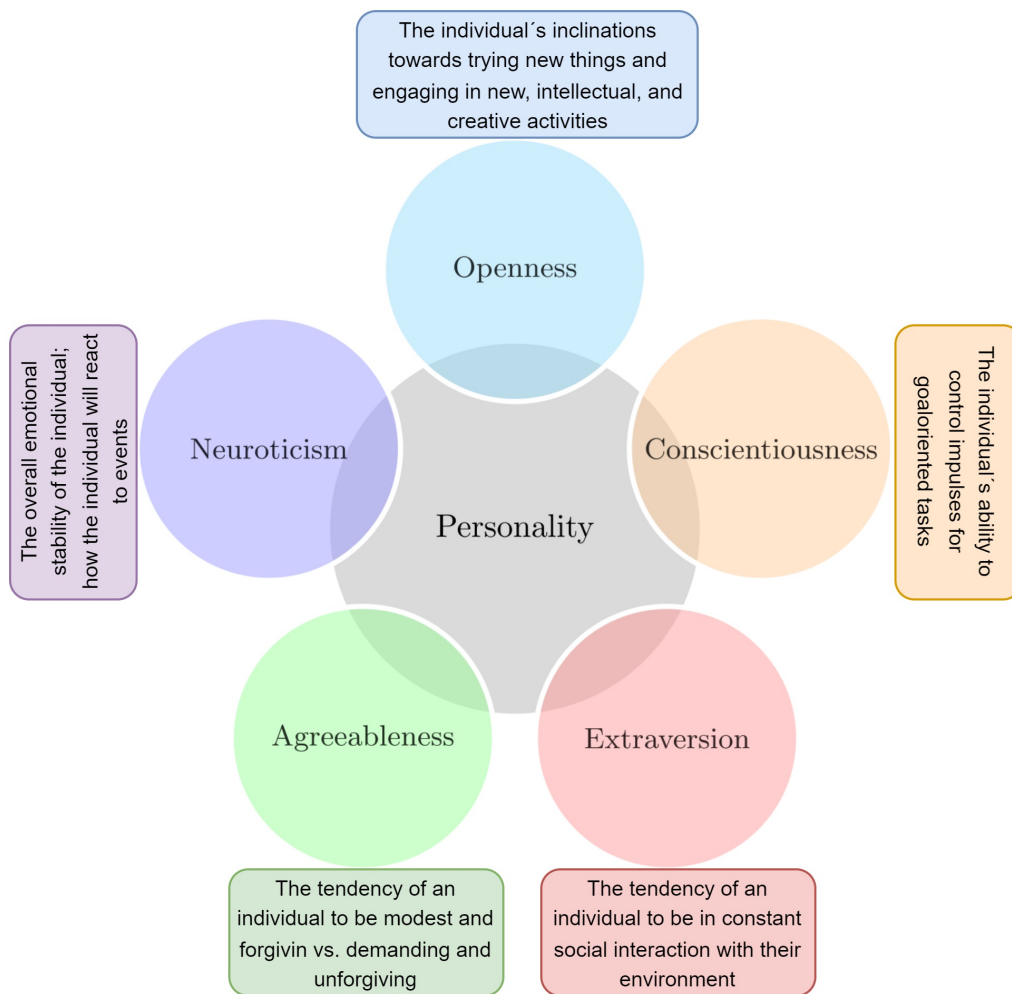


Figure 4.11: The Big Five Model (BFQ)

In addition to the literature review supporting the use of the BFQ (Big Five Questionnaire) for measuring personality in cybercriminals, it is noteworthy that these individuals often display above-average intelligence, reinforcing the appropriateness of this tool. Hackers typically exhibit openness to experience, characterized by a deep intellectual curiosity and a willingness to explore and implement cutting-edge technologies and novel attack techniques. This trait drives their innovative approaches to cybersecurity challenges. On the other hand, conscientiousness may be less straightforward; while some hackers might engage in activities that defy established norms, others demonstrate significant ethical responsibility by using their skills to uncover and address security vulnerabilities, thereby contributing positively to the field despite their unconventional methods. Therefore, the BFQ tool fits perfectly in this context, to analyze the responsibility of a hacker's behaviors or degree of openness to new learning experiences.



## 4.5. Application of the BFQ Survey

To carry out this analysis, the survey deployment was divided into three phases:

**First Phase:** The link to the form with the respective questions was published in the same forums and hacker communities where the IP address of the honeypot had been previously published. This publication included an invitation to participate in the survey for research purposes, guaranteeing the confidentiality of the data provided.

**Phase Two:** A list of known hackers was compiled, and a direct invitation to participate in the research project was sent to them. This list was based on data obtained from the honeypot and other cyber intelligence resources, ensuring that individuals relevant to the study were reached.

**Third Phase:** Once the survey responses were obtained, the triangulation technique correlated the results between the data sets. This included comparing the survey responses with the attack data recorded by the honeypot and the personality profiles. This process allowed us to identify consistent patterns and validate the findings.

Figure 4.12 illustrates how this data was represented, highlighting the correlations between the fields used in the surveys and the prevalent data in both the attack and personality datasets. This robust and detailed methodology improves the accuracy of cybercriminal profiling and provides a solid foundation for developing more effective and personalized security strategies.

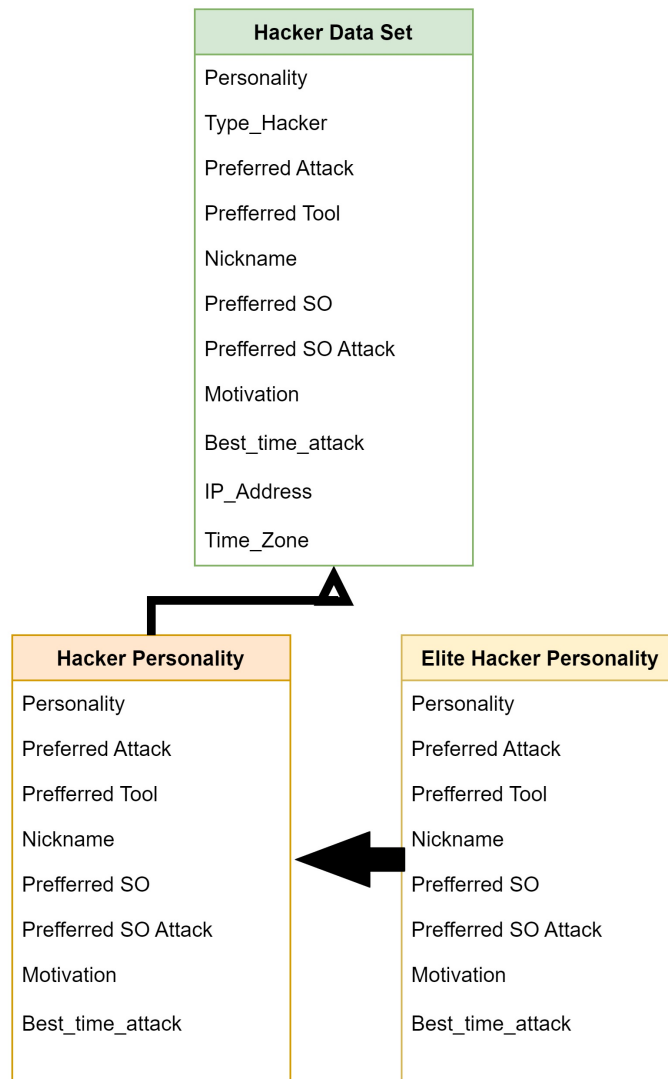


Figure 4.12: Personality and attack data set

A survey of elite hackers was conducted to collect the necessary data. With the help of the thesis director and tutor, a list of known elite hackers was compiled. Subsequently, messages were sent to the contacts found on the web, as shown in Figure 4.13.



Figure 4.13: Invite known elite hackers

Figure 4.14 shows the format of the survey sent to the hackers. This survey yielded five initial responses, which were tabulated, and used to calculate each participant's personality type. In addition to assessing personality traits, additional questions were included about the preferred type of attack, nickname used, and favorite tool, among other relevant aspects.

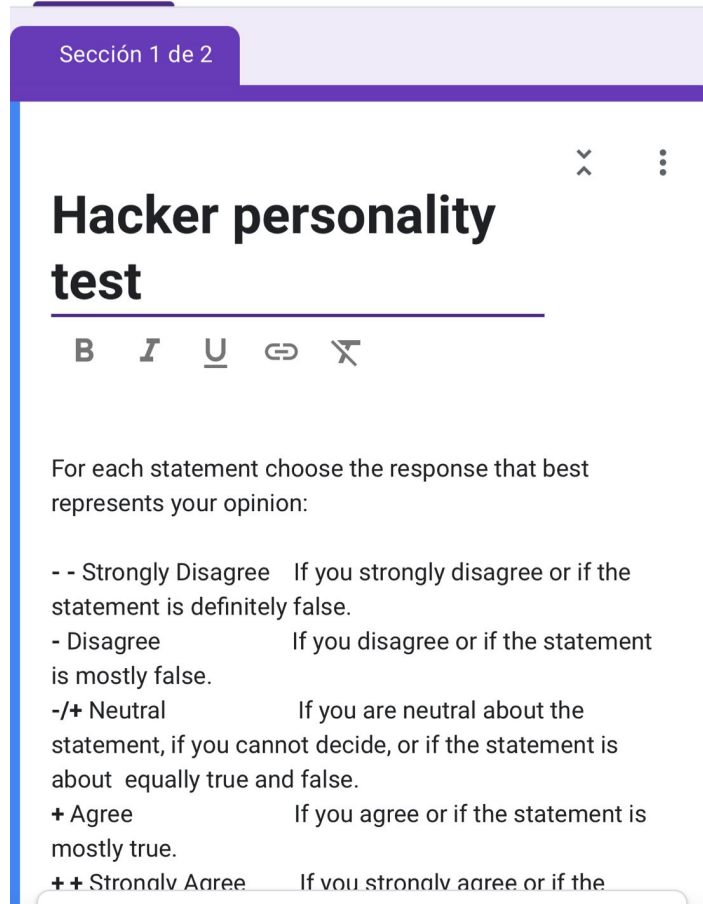


Figure 4.14: BFQ Hacker Elite Application

Simultaneously, after the survey was published in hacker forums and communities, 42 additional responses were received, as shown in Figure 4.15. After a thorough analysis, it was observed that the data obtained were consistent and similar, which validated the authenticity and reliability of the information collected for the study.

## Hacker personality test

Preguntas

Respuestas

42

Configuración

# Hacker personality test

**B** *I* U  

For each statement choose the response that best represents your opinion:

- - Strongly Disagree    If you strongly disagree or if the statement is definitely false.
- Disagree                If you disagree or if the statement is mostly false.
- /+ Neutral                If you are neutral about the statement, if you cannot decide, or if the statement is about equally true and false.
- + Agree                    If you agree or if the statement is mostly true.
- + + Strongly Agree        If you strongly agree or if the statement is definitely true.

Figure 4.15: BFQ Hackers Application

The methodology employed not only strengthened the study's validity but also provided a solid foundation for future research in the field of cybersecurity. These findings are crucial for developing more effective strategies to prevent and mitigate cyberattacks, leveraging in-depth knowledge of hackers' behavior patterns and preferences. The combination of these data allowed for a detailed and precise analysis of cybercriminal profiles, as detailed below:

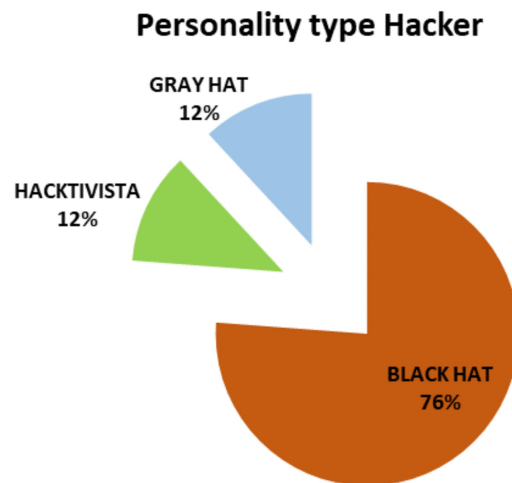


Figure 4.16: Result about the type of hacker

The majority of hackers are identified as Black Hat, with 76% of the 42 responses out of 100% falling into this category, and 12% of those who consider themselves Gray Hat and Hacktivists, as shown in Figure 4.16.

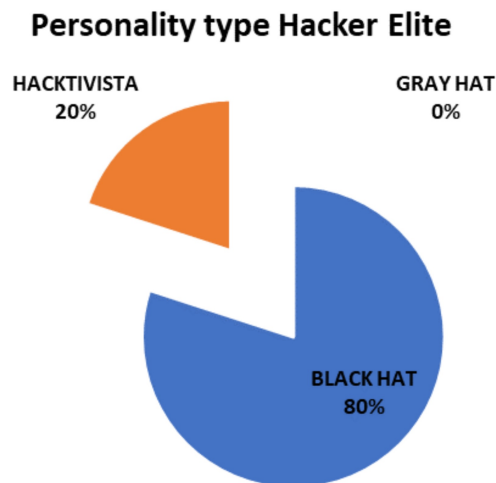


Figure 4.17: Findings about the type of elite hacker

Regarding elite hackers, 80% are Black Hat, 20% consider themselves hacktivists, and there are no values for Gray Hat, as shown in Figure 4.17.

### Hacker Personality Survey

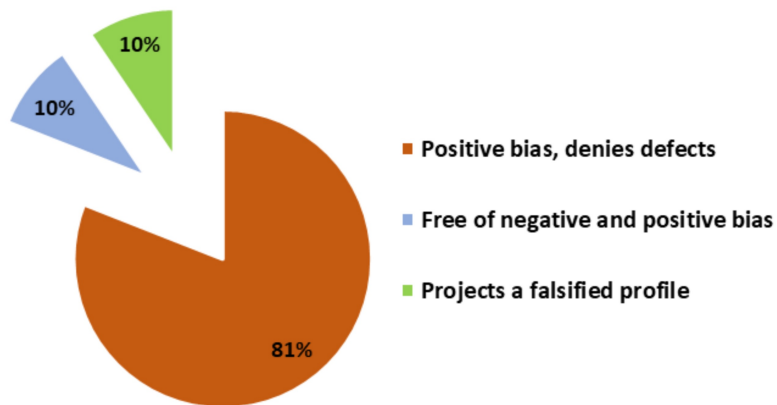


Figure 4.18: Findings on the personality of hackers

In personality assessment, "positive bias" in Big Five questionnaires refers to the tendency to perceive or interpret events, traits, or situations in an overly favorable light. Hackers may exhibit this bias by emphasizing their strengths and successes while downplaying or ignoring their weaknesses, whether in interpersonal relationships, professional settings, or their self-image. Conversely, denial of shortcomings involves an unwillingness or inability to acknowledge personal weaknesses or areas needing improvement. This denial can manifest as rejecting constructive criticism, justifying or rationalizing problematic behavior, or failing to recognize and address one's own flaws.

In the present study, as shown in Figure 4.18, we find that 81% of respondents exhibit these personality traits. Ten percent demonstrate that they are free of prejudice, with a tendency to perceive or interpret events, situations, or characteristics positively. The remaining 10% project a false profile; that is, people who responded this way do not accurately reflect their true personality; in other words, they are trying to present an image of themselves that does not correspond to how they feel, think, or behave.

### Hacker Personality Survey (Elite)

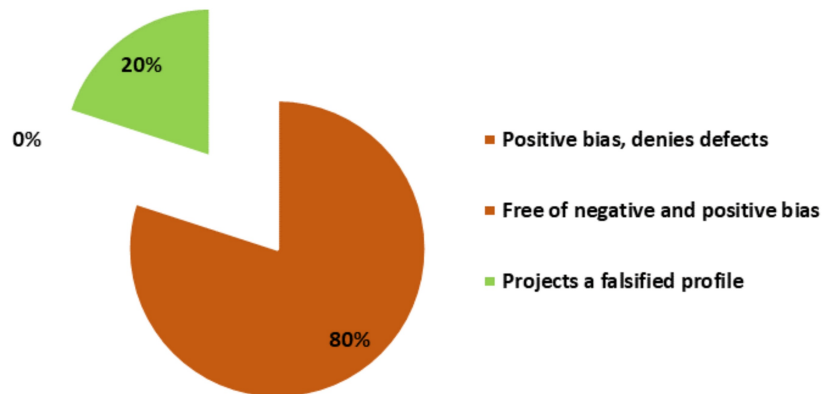


Figure 4.19: Findings on the personality of elite hackers

In the results for elite hackers, as shown in Figure 4.19, we have 80% positive bias, which refers to a tendency or inclination to perceive or interpret events, situations, or traits positively, and 20% responding in a way that does not accurately reflect their true personality.

Subsequently, machine learning (ML) techniques were employed to analyze hacking patterns, requiring data normalization into a format compatible with Python. The parameters used for analysis included Tool Attack, IP Address, Country, and Time. After processing the data, a neural network architecture was implemented. The dataset was divided into training and testing subsets, with 80% allocated for training and 20% for testing. To enhance the model's performance, the Random Forest algorithm with 50 predictors was applied. Additionally, correlation patterns between personality tests and cyberattacks were evaluated using the following fields: Personality, Train, Motivation, Country, Tool, and Time.



## 5. EVALUATION

### Contents

---

5.1 Dataset and Machine learning . . . . .	71
5.2 Discussion . . . . .	87
5.3 Contributions of this work . . . . .	89

---

### 5.1. Dataset and Machine learning

For this process, various data were used as parameters, including attack tools, IP addresses, countries of origin, times of the attacks, and types of attacks. These parameters were carefully selected to view cyber attackers' behavior comprehensively. Table 5.1 presents a detailed classification of these parameters, showing how they are grouped and analyzed to obtain valuable information. This approach allowed for the identification of patterns and trends that are essential for better understanding the tactics employed by cybercriminals and strengthening defense strategies.

---

<b>Parameter</b>	<b>type of parameters</b>
Tool Attack	Categorical or class
IP address	Integer
Country	Categorical or class
Time	Integer

---

Table 5.1: Research parameters

The parameter to be classified is based on the type of attack, which is also considered a category or class. A histogram of the existing data is made for categorical data, and the files are stored with the unique names in each field. For example, from the database for the Country parameter, the following is obtained, as shown in Figure 5.1:

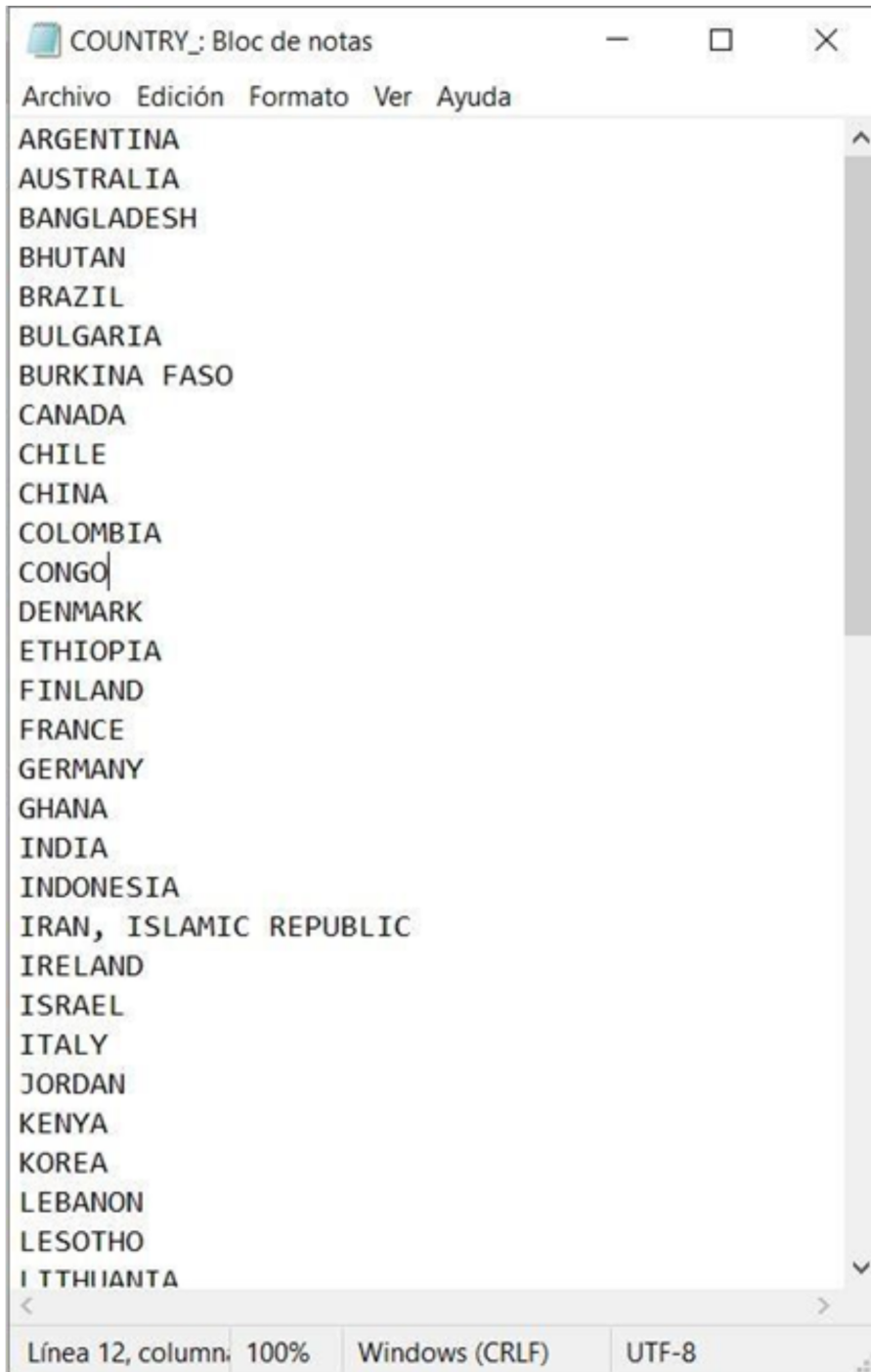


Figure 5.1: Classifier parameter contry

The parameter related to the IP address had to be converted to an integer value to manage the data during classification and the parameter time in hours, minutes, and seconds. In addition, according to the data, there are seven different types of attacks. According to the values taken by the IP Address, Country, Time, and Attack Tool, the system

should be able to predict what type of attack the server may suffer. The data are extracted from files with CSV extension, previously saved for each type using the PANDAS library in Python, as can be seen in the following link: <https://colab.research.google.com/drive/1bz2YK1MJOZrSRXBQb3h-wdFaXxiEfl?usp=sharing> .

At this point, the performance metrics were defined to evaluate whether the machine learning algorithm implemented in this research is a neural or artificial neural network (ANN). The confusion matrix is composed of prediction data (non-hacked data) and real data (hacked data); in this table for its classification, logical values of 0 as negative and one as positive were sought to facilitate the language of values a "No and a Yes." Each of these applications is assigned a defined vector, which contains the information on the permissions and the classification labels, as shown in Table 5.2.

multirow

	<b>REAL</b>	
<b>PREDICTION</b>	DP FP	FN DN

Table 5.2: Structure of the confusion matrix

False positives (FP) represent a yes but, in reality, a no.

False negative (FN) values represent a no, but it is a yes.

Positive data (PD) values that represent a yes.

Negative data (ND) values that represent a no.

Four evaluation metrics Loss, Accuracy, Recall, Accuracy, and F1 Score, were used to evaluate the classification performance of the problem, which are defined as follows:

Loss is a penalty for misclassification. The Loss function to be used is binary cross-entropy, defined by the following formula.

$$y_{i,l} \in \{0, 1\} \wedge l \in [1, L] \wedge i \in [1, N] \quad (5.1)$$

$$binarycrossentropy = -y_{i,l} \cdot \ln(\hat{y}_{i,l}) + (1 - y_{i,l}) \cdot \ln(1 - \hat{y}_{i,l}) \quad (5.2)$$

Accuracy: this metric is calculated from the number of correctly classified values.

$$Accuracy = \frac{DP + DN}{DP + DN + FP + FN} \quad (5.3)$$

The recall metric is calculated from correctly predicted positive data over the total positive data plus the test set.

$$Recall = \frac{DP}{DP + FN} \quad (5.4)$$

Accuracy is calculated from correctly predicted positive data over total predicted positive data.

$$Precision = \frac{DP}{DP + FP} \quad (5.5)$$

F1 Score is interpreted as the harmonic measure between accuracy and Recall, where F1 shows the best and worst scores.

$$F1Score = 2 + \frac{Precision * Recall}{Precision + Recall} \quad (5.6)$$

To improve the accuracy, Random Forests will be used, and its representation is usually

$$f_t(x) = f(x, \theta) \quad (5.7)$$

the form denotes the whole forest

$$F = f_1 \dots \dots \dots f_T \quad (5.8)$$

Where T is the number of trees in the forest, with the following formula representing the probability of prediction of class k

$$p_{k/x} = \frac{1}{T} \sum_{t=1}^T p_t^{k/x} \quad (5.9)$$

Where

$$p_t^{k/x}$$

is the estimated density of data classification levels. The final functional is defined as

$$C(x) = \arg \max P_{k/x} k \in \Upsilon \quad (5.10)$$

As mentioned above, to develop the neural network architecture, it was necessary to divide the data into four sets with the criterion of 80% of the data for training and 20% for testing. Also, the TensorFlow library was used; for this case, the input layer must have four neurons, one for each parameter, and seven neurons in the output layer; each one will identify a type of attack; this model has two hidden layers to improve the results. The network architecture, made using the playground-TensorFlow tool, is presented in Figure 5.2.

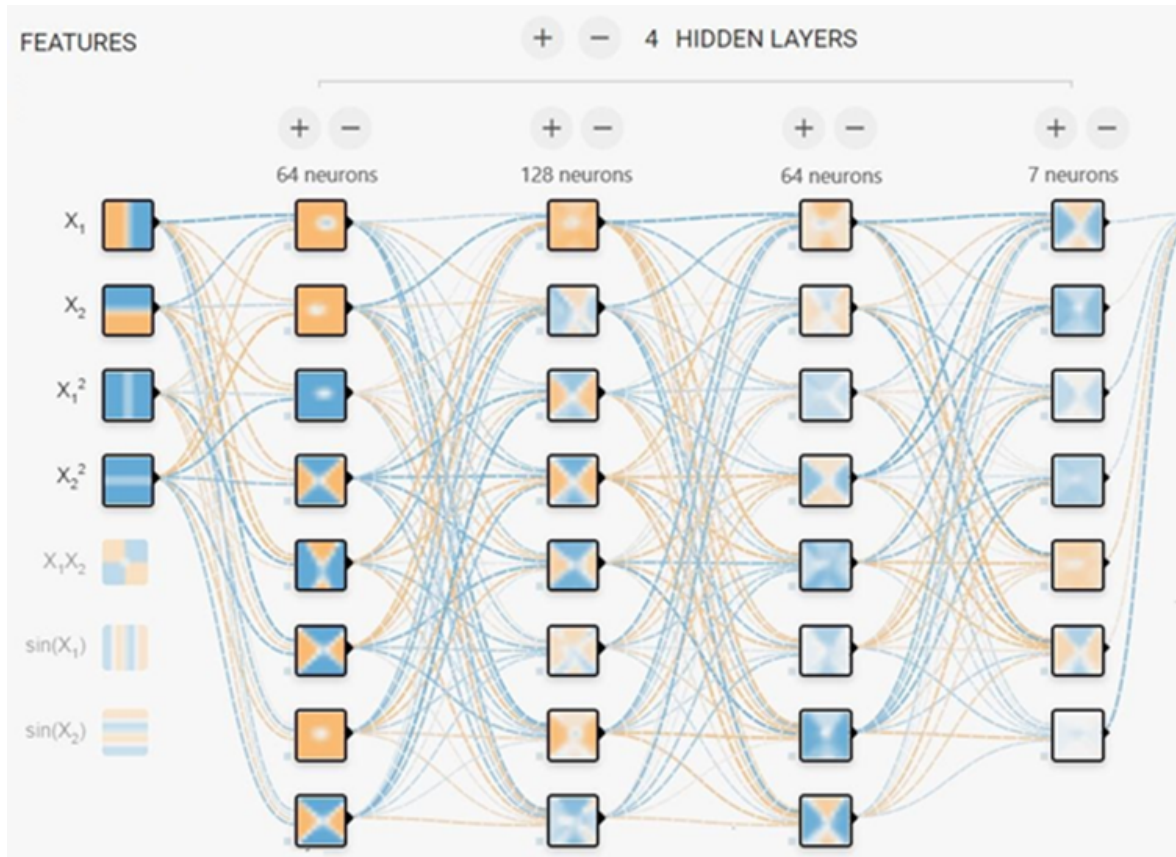


Figure 5.2: Architecture of the Neuronal Network using a Sequential Model

On the other hand, the optimized RMSPROP was selected, the error known as the LOSS metric using categorical cross-entropy, and finally, the metric to be evaluated is the accuracy (ACCURACY). The training process ran 100 epochs with a Batch of 16 data for each epoch; as the epochs pass, it is expected that both the training error and the validation will decrease. To evaluate the model fit of the data, the confusion matrix was used to estimate the error at the time of classification using the SEABORN library, as shown in Table 5.3 and Figure 5.3.

	Precision	Recall	F1-score	Support
<b>BRUTE FORCE ATTACKS</b>	0.93	1.00	0.96	41
<b>COMMAND INJECTION</b>	1.00	1.00	1.00	31.00
<b>CROSS-SITE SCRIPTING (XSS)</b>	1.00	1.00	1.00	47.00
<b>DDoS ATTACK</b>	1.00	1.00	1.00	157.00
<b>DEFACEMENT</b>	1.00	1.00	1.00	25.00
<b>SQL INJECTION</b>	1.00	1.00	1.00	79.00
<b>THE MAN IN THE MIDDLE</b>	1.00	0.85	0.92	20.00
<b>Accuracy</b>			0.99	400
<b>Macro avg</b>	0.99	0.98	0.98	400
<b>Weighted avg</b>	0.99	0.99	0.99	400

Table 5.3: Classification report of the training data the model already knew

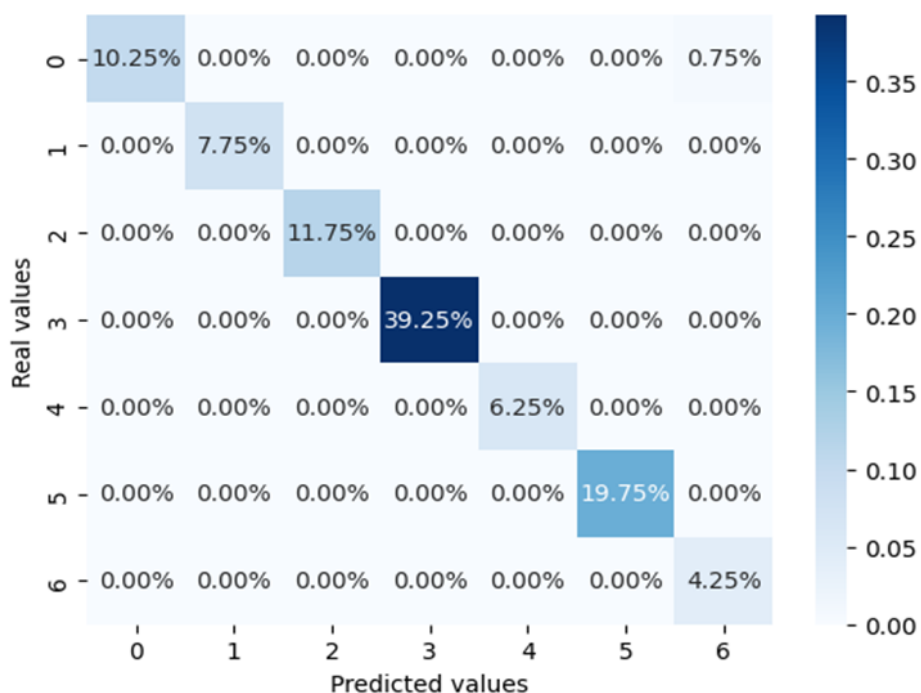


Figure 5.3: Confusion Matrix with the training data that the model already knew

After that, the results generally align with expectations, but class four has a problem. One way to evaluate the model's fit to the data is to pay attention to the distribution of the observations using the density estimation by a KDE Kernel, as shown in the following plot in

Figure 5.4 and Table 5.4.

	Precision	Recall	F1-score	Support
<b>BRUTE FORCE ATTACKS</b>	0.91	1.00	0.95	10
<b>COMMAND INJECTION</b>	1.00	1.00	1.00	5.00
<b>CROSS-SITE SCRIPTING (XSS)</b>	1.00	1.00	1.00	12.00
<b>DDOS ATTACK</b>	1.00	1.00	1.00	38.00
<b>DEFACEMENT</b>	0.67	1.00	0.80	2.00
<b>SQL INJECTION</b>	1.00	1.00	1.00	25.00
<b>THE MAN IN THE MIDDLE</b>	1.00	0.75	0.86	8.00
<b>Accuracy</b>			0.98	100
<b>Macro avg</b>	0.94	0.96	0.94	100
<b>Weighted avg</b>	0.98	0.98	0.98	100

Table 5.4: Classification report with training data not known to the model

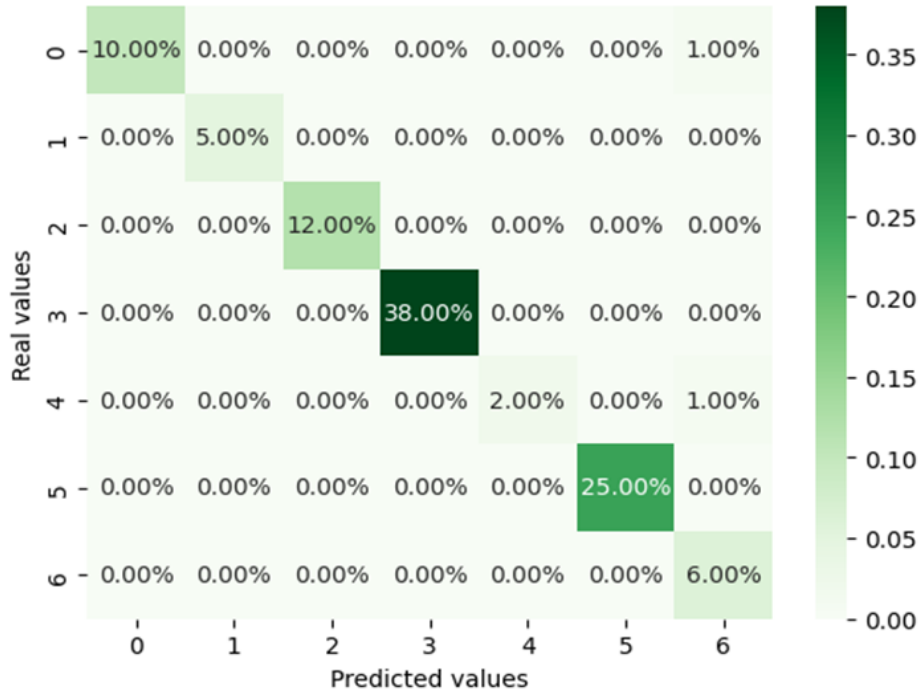


Figure 5.4: Confusion Matrix with training data not known to the model

Figure 5.5 shows the data, illustrating the density of the observations predicted by the model to the actual observations. This graph provides a visual representation of how the

model predicts the distribution of different types of attacks compared to the real data, allowing for a detailed assessment of its accuracy and effectiveness. Presenting these data is essential for understanding the model's ability to accurately capture the trends and patterns of attacks, thereby reinforcing the model's validity and utility in practical cybersecurity applications.

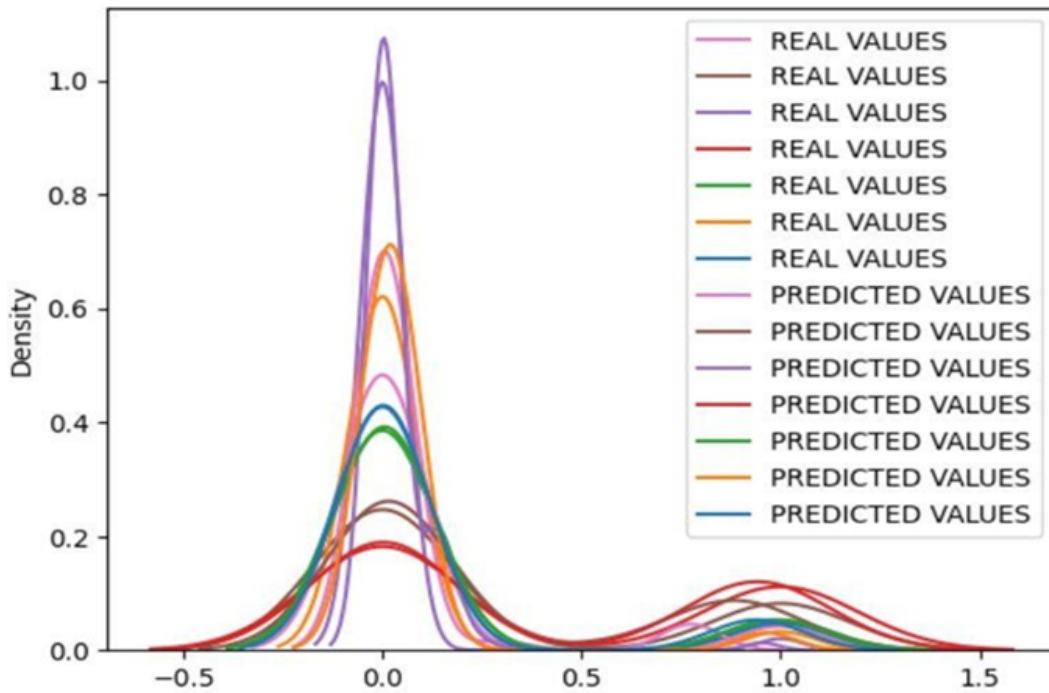


Figure 5.5: Density of model-predicted observations matched to actual observations

To improve the results obtained with the previous architecture, we used the RANDOM FOREST algorithm. The data were normalized and divided into training and test sets. Moreover, it was used with 50 predictors for this problem. The prediction with the already known set is shown in the following chart. In Figure 5.6 and Table 5.5, the following data are found in the matrix, demonstrating that the model correctly differentiates all classes of attacks with known data. This confusion matrix illustrates the model's accuracy in classifying each type of attack, highlighting its ability to identify various categories of cyber threats correctly. Presenting these data is crucial to validate the model's effectiveness in detecting and classifying cyberattacks, confirming its reliability and robustness in a real-world environment.



	Precision	Recall	F1-score	Support
<b>BRUTE FORCE ATTACKS</b>	1.00	1.00	1.00	10
<b>COMMAND INJECTION</b>	1.00	1.00	1.00	5.00
<b>CROSS-SITE SCRIPTING (XSS)</b>	1.00	1.00	1.00	12.00
<b>DDOS ATTACK</b>	1.00	1.00	1.00	38.00
<b>DEFACEMENT</b>	1.00	1.00	0.80	2.00
<b>SQL INJECTION</b>	1.00	1.00	1.00	25.00
<b>THE MAN IN THE MIDDLE</b>	1.00	1.00	1.00	8.00
<b>Accuracy</b>			1.00	400
<b>Macro avg</b>	1.00	1.00	1.00	400
<b>Weighted avg</b>	1.00	1.00	1.00	400

Table 5.5: Classification report the model differentiates all attack classes correctly with known data

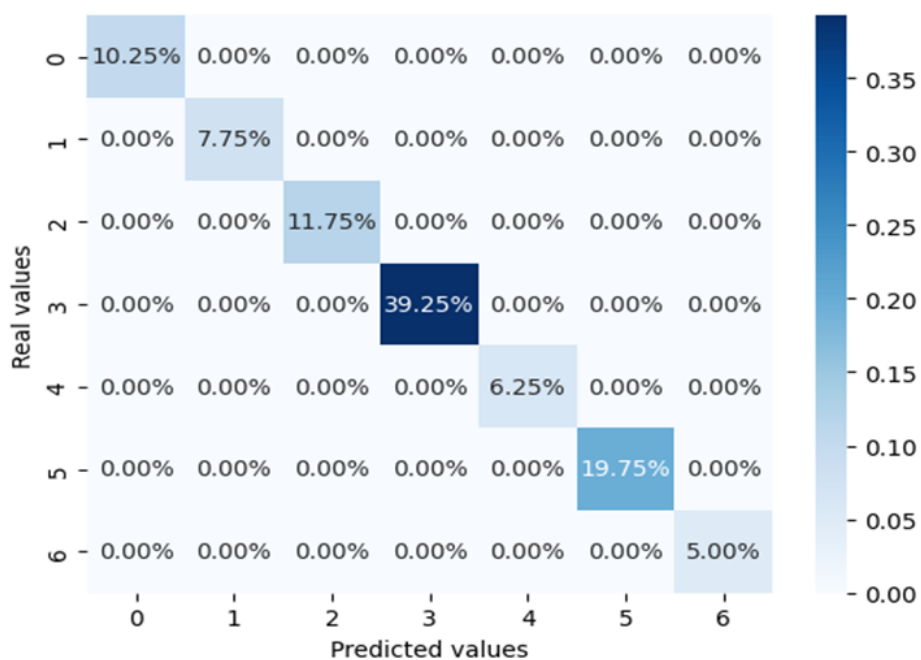


Figure 5.6: Confusion Matrix of the model differentiates all attack classes correctly with known data

In Figure 5.7 and Table 5.6, a matrix shows that the model correctly distinguishes all classes of attacks using unknown data. This visual representation is crucial to demonstrate

the model's ability to accurately classify each type of attack, even when confronted with new or previously unlabeled data. Accuracy in this task is essential for enhancing confidence in the model and its practical application in detecting and mitigating cyber threats.

	<b>Precision</b>	<b>Recall</b>	<b>F1-score</b>	<b>Support</b>
<b>BRUTE FORCE ATTACKS</b>	0.91	1.00	0.95	10
<b>COMMAND INJECTION</b>	1.00	1.00	1.00	5.00
<b>CROSS-SITE SCRIPTING (XSS)</b>	1.00	1.00	1.00	12.00
<b>DDOS ATTACK</b>	1.00	1.00	1.00	38.00
<b>DEFACEMENT</b>	0.67	1.00	0.80	2.00
<b>SQL INJECTION</b>	1.00	1.00	1.00	25.00
<b>THE MAN IN THE MIDDLE</b>	1.00	0.75	0.86	8.00
<b>Accuracy</b>			0.98	100
<b>Macro avg</b>	0.94	0,96	0.94	100
<b>Weighted avg</b>	0.98	0.98	0.98	100

Table 5.6: Classification report of model differentiates all attack classes correctly with unknown data

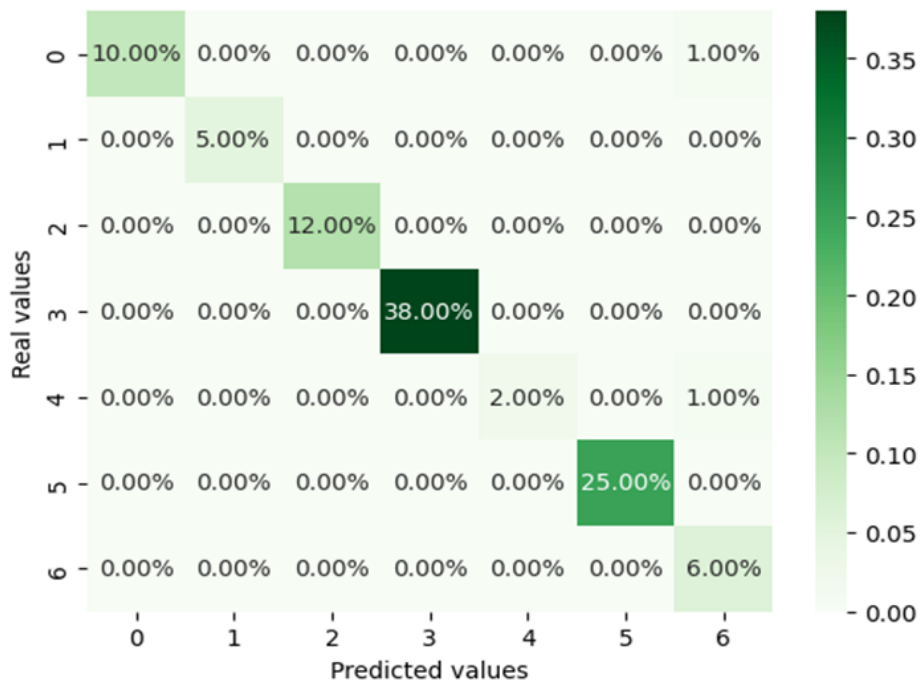


Figure 5.7: Confusion Matrix of model differentiates all attack classes correctly with unknown data

The system has improved the previously obtained results, although it still maintains errors when differentiating class four when faced with unknown data. In this context, the decision tree is structured to prioritize certain key parameters in data classification, as indicated by Figure 5.8, which is structured as follows:

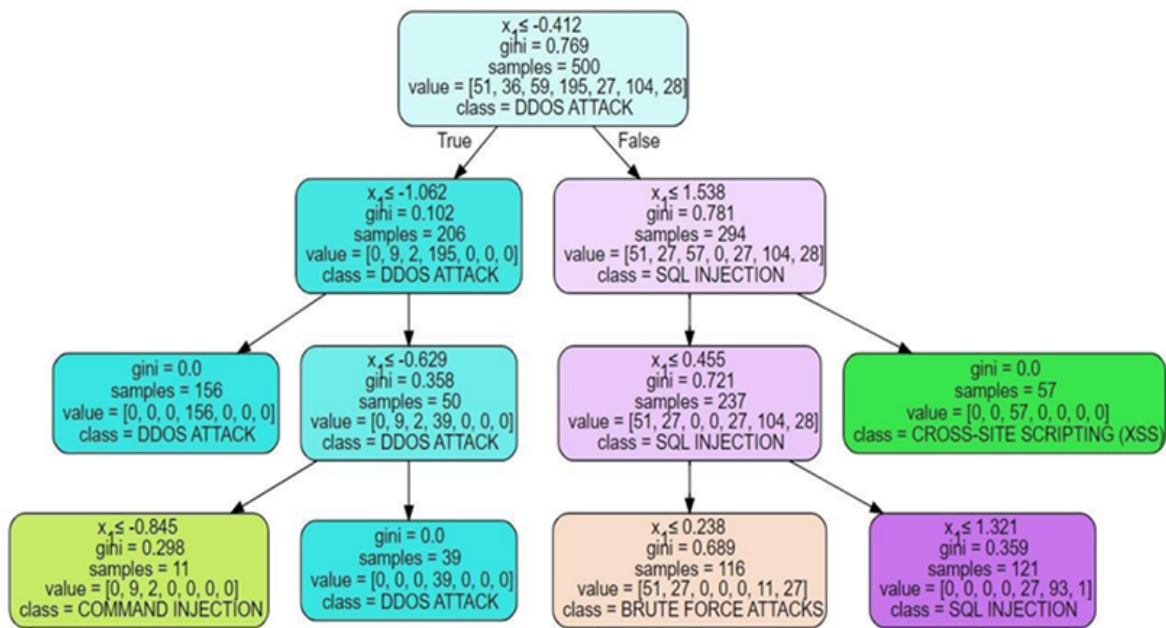


Figure 5.8: Classification decision tree

The previously described procedures were applied to verify the relationship between the personality of hackers or cybercriminals and the attacks they perpetrate. These procedures include collecting and analyzing data through surveys and activity logs obtained via honeypots. The fields used for this analysis are detailed in Table 5.7. This table provides a comprehensive classification of the evaluated parameters, ranging from personality traits to types of attacks, tools used, and other relevant factors. This approach allows for a deeper understanding of how psychological characteristics can influence the tactics and behaviors of cybercriminals, providing a solid foundation for developing more effective and personalized security strategies.

Parameter	type of parameters
Personality	Categorical or class
Trait	Categorical or class
Motivation	Categorical or class
Country	Categorical or class
Tool	Categorical or class
Time	Integer

Table 5.7: Type of parameters

Similarly, the corresponding histograms for each parameter were generated. Once the data were loaded, the correlation matrix between the indicated parameters was evaluated. The evaluation revealed the presence of multicollinearity among the parameters, indicating significant interdependencies between them. This analysis used six specific parameters for the case study, and 250 predictors defined the model. Figure 5.9 shows the obtained results, visually representing the correlations and the data distribution. This figure is crucial for understanding the dataset's internal relationships and adjusting the model more precisely, thereby optimizing the predictive capacity and robustness of the analytical system.

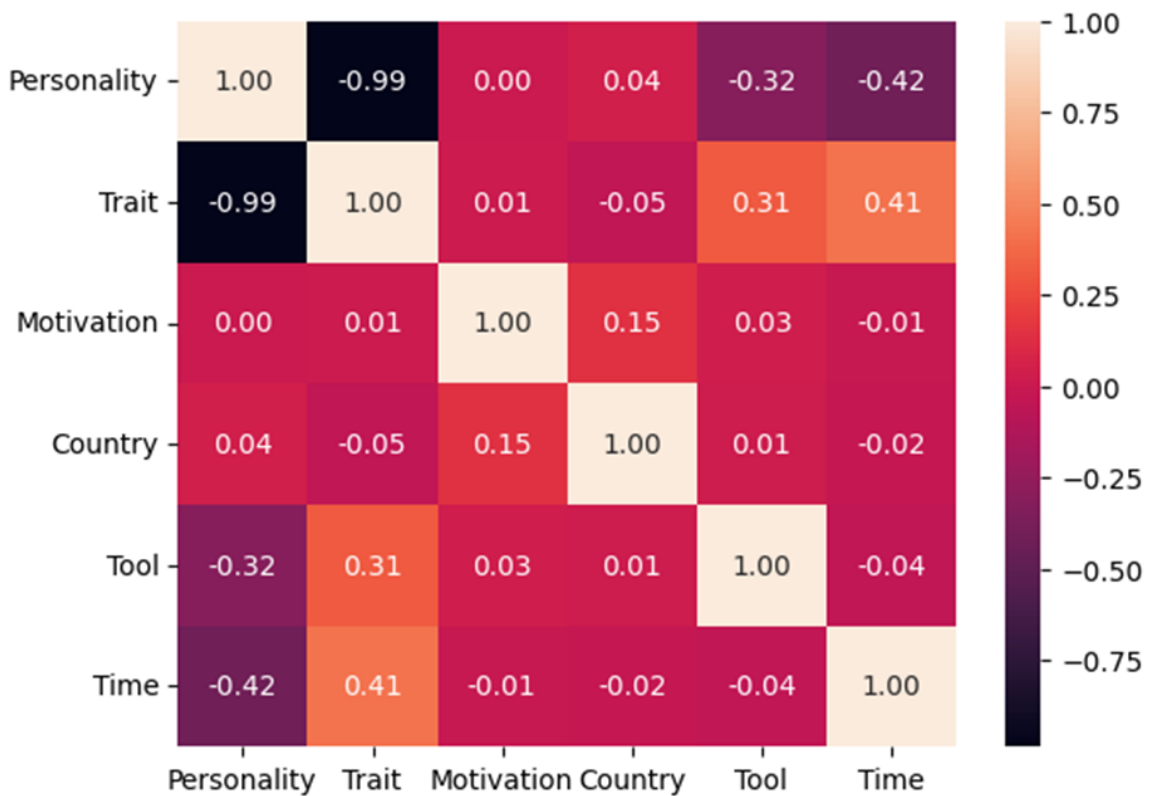


Figure 5.9: Multi-correlation between hackers personality data

In this case, the model produces expected values that allow for the classification of the seven types of attacks defined in this study. This classification is based on certain personality traits and additional characteristics such as the country of origin and the timing of the attack. Figure 5.10 and Table 5.8 illustrates how these factors influence the model's accuracy in distinguishing between different types of cyberattacks. Integrating these parameters improves the model's accuracy and provides a deeper understanding of attackers behavior patterns and trends. These findings are essential for developing more effective and person-

alized cybersecurity strategies capable of proactively anticipating and mitigating threats.

	Precision	Recall	F1-score	Support
<b>BRUTE FORCE ATTACKS</b>	1.00	1.00	1.00	41
<b>COMMAND INJECTION</b>	1.00	1.00	1.00	31.00
<b>CROSS-SITE SCRIPTING (XSS)</b>	1.00	1.00	1.00	47.00
<b>DDOS ATTACK</b>	1.00	1.00	1.00	157.00
<b>DEFACEMENT</b>	1.00	1.00	1.00	25.00
<b>SQL INJECTION</b>	1.00	1.00	1.00	79.00
<b>THE MAN IN THE MIDDLE</b>	1.00	1.00	1.00	20.00
<b>Accuracy</b>			1.00	400
<b>Macro avg</b>	1.00	1.00	1.00	400
<b>Weighted avg</b>	1.00	1.00	1.00	400

Table 5.8: Classification report with the personality test data known to the model

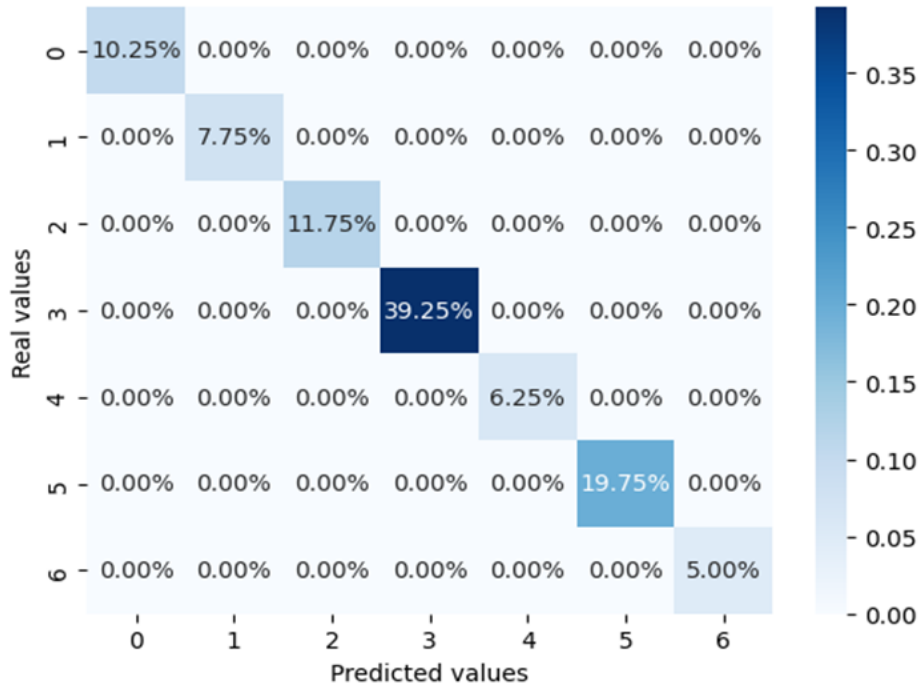


Figure 5.10: Confusion Matrix with the personality test data known to the model

When evaluating the model, a distribution of the observations is observed by estimating the density using Kernel Density Estimation (KDE). This method allows the data to be

correctly adjusted with the predictions of the Machine Learning (ML) model, providing a clear and accurate visual representation of how the model performs about the actual data. Figure 5.11 and Table 5.9 illustrates this fit, showing how the KDE density estimates align with the model's predictions. This analysis is crucial for validating the model's effectiveness and ensuring that the predictions are consistent with real observations, which enhances the model's reliability in practical cybersecurity applications.

	<b>Precision</b>	<b>Recall</b>	<b>F1-score</b>	<b>Support</b>
<b>BRUTE FORCE ATTACKS</b>	0.90	0.82	0.86	11.00
<b>COMMAND INJECTION</b>	0.80	0.80	0.80	5.00
<b>CROSS-SITE SCRIPTING (XSS)</b>	1.00	0.92	0.96	13.00
<b>DDOS ATTACK</b>	1.00	1.00	1.00	38.00
<b>DEFACEMENT</b>	0.50	0.50	0.50	2.00
<b>SQL INJECTION</b>	0.96	0.92	0.40	26.00
<b>THE MAN IN THE MIDDLE</b>	0.66	1.00	0.77	5.00
<b>Accuracy</b>			1.00	100
<b>Macro avg</b>	1.00	1.00	1.00	100
<b>Weighted avg</b>	1.00	1.00	1.00	100

Table 5.9: Classification report with the personality test data that the model does not know



Figure 5.11: Confusion Matrix with the personality test data that the model does not know

Figure 5.12, titled "Density of observations predicted by the model matched to the actual observations of the hacker patterns," provides a detailed view of how the model's predictions align with the observed real data. This figure highlights the model's accuracy in identifying and replicating hacker behavior patterns, showing that the densities estimated by the model closely match the real observations. This level of precision is essential to validate the model's ability to correctly predict cyber attackers behaviors, thereby providing a solid foundation for developing more effective and proactive defense strategies in the field of cybersecurity.



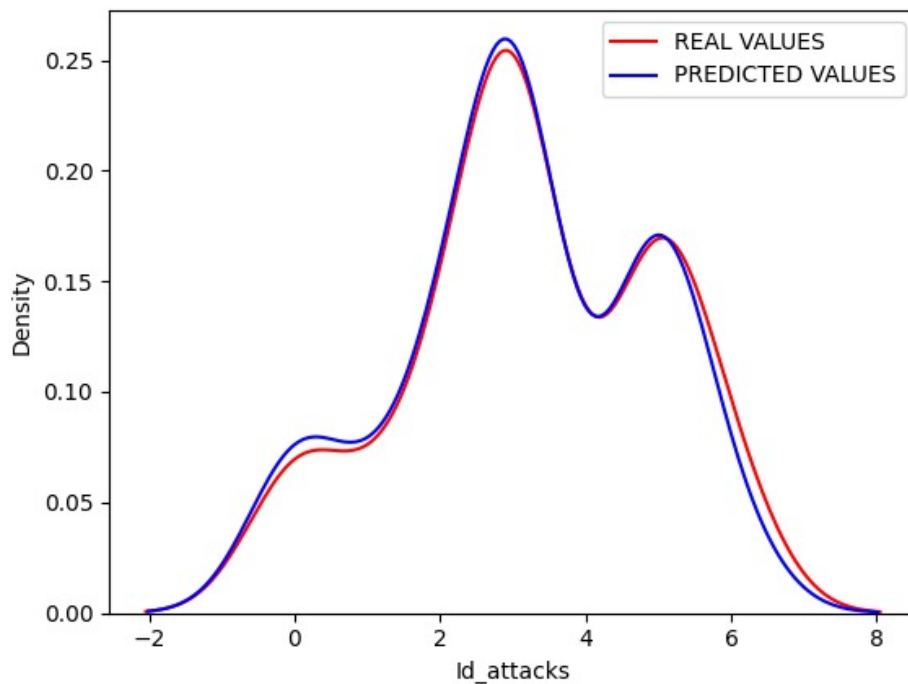


Figure 5.12: Density of observations predicted by the model matched to the actual observations of the hacker patterns

## 5.2. Discussion

This research adopts a distinctive approach by integrating and comparing three critical areas: honeypots, machine learning for cyber-attack detection, and the analysis of cyber-attack personality traits. Traditionally, these elements have been studied in isolation, but this research seeks to bridge these gaps, offering a novel perspective in cybersecurity. By synthesizing these areas, the study aims to provide groundbreaking insights into cyber-attack patterns and defensive strategies

The existing literature highlights the effectiveness of honeypots in capturing and analyzing attackers tactics, techniques, and procedures. Designed to attract and capture cyber-criminals, honeypots offer a rich data source for understanding malicious behaviors. Their strategic use in cybersecurity research has proven invaluable, shedding light on the complex dynamics of cyber-attacks and enhancing our ability to devise effective countermeasures. Studies such as *The Use of Honeypot in Machine Learning Based on Malware Detection: A Review* [112] highlight the synthesis of various studies and approaches that have integrated honeypots and machine learning, underscoring the benefits and challenges of these meth-

ods. Moreover, this research identifies opportunities to enhance the effectiveness of these technologies and proposes best practices for their implementation. The findings demonstrate that integrating honeypots with machine learning is both highly effective and practical for real-time attack detection. This is due to the algorithms capacity to identify subtle and unusual patterns in the data, as evidenced in our study. The practical implications of our results underscore the relevance and applicability of our research in addressing real-world cybersecurity challenges.

The study Predicting Attack Patterns via Machine Learning by Exploiting Stateful Firewall as Virtual Network Function in an SDN Network [116] offers several notable contributions to enhancing the security of software defined networks (SDNs). By integrating stateful firewalls as virtual network functions (VNFs) and using machine learning techniques to analyze network traffic data collected by these firewalls, the research significantly advances attack detection and prevention. This approach achieves high accuracy in attack prediction, improves threat detection performance, and reduces false positives. Our proposal emphasizes attack signatures and classifies the cognitive patterns of hackers, providing practical and effective methods for cyber-attack detection and prevention.

The article Enhancing Network Security through a Multi-layered Honeypot Architecture with Integrated Network Monitoring [123] presents a multi-layered honeypot architecture integrated with advanced monitoring tools to enhance network security by effectively collecting and analyzing attack data. Our research differs significantly by combining this data with personality trait analysis to provide a deeper and more holistic understanding of cyber attackers.

These studies have shown that machine learning can improve the accuracy and efficiency of attack detection. Our research extends this approach by using machine learning to detect attack patterns and analyze and predict personality traits based on behavior observed through honeypots and survey data. For example, we have identified and analyzed cyberattackers personality traits in personality trait analysis by combining data from honeypots and surveys. Previous studies such as ¿Is there a cybercriminal personality? Comparing cyber offenders and offline offenders on HEXACO personality domains and their underlying facets [143] attempt to determine whether there are significant differences in personality traits between cybercriminals and non-cybercriminals using the HEXACO personality model. This model assesses six major domains: Honesty-Humility, Emotionality,

Extraversion, Agreeableness, Conscientiousness, and Openness to Experience. It was applied to 200 cybercriminals, 200 non-cybercriminals, and 200 individuals with no criminal record. The results show that cybercriminals have a distinctive personality profile compared to non-cybercriminals and individuals with no criminal record. These differences may be critical for developing more effective prevention strategies and rehabilitation programs tailored to the unique characteristics of cybercriminals.

The study Psychological Profiling of Hacking Potential [136] aims to identify and analyze psychological and personality traits that may predict an individual's likelihood of engaging in hacking activities. By combining psychological and criminological perspectives, the research seeks to deepen the understanding of the factors motivating hackers and the traits that differentiate them from the broader population. The study's sample of 500 individuals was meticulously selected to encompass a range of interests and levels of involvement in hacking, ensuring a robust analysis. The findings offer a nuanced view of the personality traits and motivational factors that could indicate a propensity for hacking. The literature review supports the existence of these motivational factors and highlights that each hacker typically develops a unique attack strategy.

### **5.3. Contributions of this work**

In today's digital era, information security has become a critical concern for individuals, businesses, and institutions globally. The increasing sophistication of cyber-attacks necessitates the development of advanced methods for detecting and preventing such threats. Machine learning models, particularly those leveraging artificial intelligence, have emerged as powerful tools in this domain. These models excel at analyzing vast amounts of data to uncover subtle patterns indicative of malicious activities. In cybersecurity, they are instrumental in detecting anomalous behavior, identifying potential threats, and preemptively thwarting attacks.

A key contribution of this study is its innovative approach, which integrates honeypots and surveys to gather data on cyber-attacks and the personality traits of cybercriminals. This combination provides invaluable insights into the tactics and techniques used by attackers, enabling the development of more precise and effective machine learning models. By incorporating technical data from honeypots with psychological profiles obtained from surveys, this research facilitates the creation of sophisticated models for detecting and preventing

cyber-attacks. These models can analyze large amounts of data to identify subtle patterns indicating malicious activity. In information security, learning models can detect anomalous behavior, identify potential threats, and prevent attacks before they cause damage.

In addition to honeypots, surveys are used to gather information about the personality traits of individuals involved in cyberattacks; these surveys can help identify common psychological profiles among attackers, which in turn can inform the development of more sophisticated learning models. By combining technical data from honeypots with psychological information from surveys, researchers can create and accurate models for detecting and preventing cyber-attacks.

Applying these learning models based on data collected through honeypots and surveys significantly impacts information security. These models can identify patterns of suspicious behavior with greater accuracy, enabling organizations to detect and respond to cyber threats quickly. By better understanding the personality traits of attackers, organizations can strengthen their defenses and develop more effective strategies to protect their data and systems. In addition, the ability to anticipate and prevent cyber-attacks can have a significant economic and reputational impact on organizations. By implementing advanced learning models, companies can reduce the risk of costly security breaches and mitigate the negative impact on their reputation.

It is important to keep in mind that cyber security combines technical aspects with an understanding of human psychology to protect systems and data from cyber threats; therefore, it is critical to implement sound technical measures, but it is also crucial to consider human aspects such as hacker motivations, personality traits, and security culture to develop effective cyber protection strategies; by integrating these elements, organizations can strengthen their security posture and mitigate the risks associated with the ever-evolving cyberspace.

In addition, software vendors should regularly issue security patches to address new threats and vulnerabilities. It should be noted that the use of secure protocols such as HTTPS (HTTP over SSL/TLS) for web communications helps protect the privacy and integrity of transmitted data; other protocols such as SSH (Secure Shell) and VPN (Virtual Private Network) also ensure secure communications and remote access. These technical aspects of IT security are essential to protect information and digital infrastructure in an increasingly networked and cyber-attack-prone environment.

IT Security is not just about implementing technical technologies and measures; it is important to recognize that IT security is not just about technology; it also involves psychological and human aspects, which is why understanding the nuances of human psychology is essential to designing effective security strategies; understanding how people perceive and respond to security policies is critical to promoting compliance, noting that factors such as convenience, risk perception, and trust in the system influence the degree to which individuals comply with security policies; finding a balance between security and usability to avoid fatigue and encourage secure behavior; noting that education and training in computer security are critical to raising awareness of risks and promoting secure behavior, which should include training in recognizing phishing attacks, good password practices, and how to protect confidential information.

Organizational culture is important in IT security, as it influences employee attitudes and behaviors toward security. Fostering a security culture prioritizing information protection will help reduce risk and improve an organization's security posture.

This research, through the analysis of personality ranges, identifies that people may experience a variety of emotions, such as anxiety, fear, or frustration, when they are victims of security incidents, so understanding these emotional reactions is important to provide adequate support and mitigate the psychological impacts of security incidents, so that these nuances of human psychology highlight the importance of considering the human factor in computer security. While technical measures are essential, addressing people's attitudes, perceptions, and behaviors is crucial to ensure effective security.

For example, hackers are often highly adaptable and deeply understand computer systems and underlying technologies. This technical understanding allows them to identify new techniques and tools to circumvent security measures, highlighting the importance of keeping up to date with the latest trends in computer security. However, these individuals are often persistent in their efforts to compromise systems, showing remarkable resilience to security obstacles and countermeasures; this persistence underscores the need to implement robust defenses and incident response measures to protect data against persistent attacks, so it should always be kept in mind that hacker personality traits have a significant impact on data security, whether through identifying vulnerabilities, exploiting systems or contributing to the improvement of IT security. Understanding these traits is crucial to devel-

oping effective data protection and risk mitigation strategies in an increasingly complex and dynamic digital environment.

"Learning Models to Detecting Personality Traits of Cyber Attackers: A Combined Approach Using Honeypot and Surveys" represents a significant contribution to information security; by harnessing the power of learning models and combining technical data with psychological information, this approach provides a more complete and accurate view of cyber threats, and by improving the ability to detect and prevent attacks, these models help protect data integrity, confidentiality, and availability in an increasingly interconnected and vulnerable digital world.

Finally, this research proves that there are common patterns among attackers or patterns a hacker follows before an attack. This generates value for this work and motivates researchers to continue studies in this field since if we know a hacker, we can predict his attack and minimize the risk. For this, we also provide a data set of patterns different from the common databases, such as the most desirable time of day for hackers depending on their location, tool used, or preferred attack. This data will allow us to make better inferences to provide more effective solutions to IT infrastructures.

## 6. CONCLUSIONS AND FUTURE WORK

### Contents

---

6.1 Conclusions . . . . .	93
6.2 Future work . . . . .	96

---

### 6.1. Conclusions

With the digital evolution of companies and public institutions nationwide, users have become perfect targets for cybercriminals, who use the web to be anonymous and carry out attacks. Institutions have increased and improved their security systems to minimize the probability of being victims. That is why ESPOCH is not out of the security detection systems, since it allowed the installation of the server infrastructure to detect attacks and thus have the knowledge and seek solutions to threats.

Web server security measures are currently provided by firewall rules or WAF devices for packet analysis. Analyzing data other than those mentioned above will improve and innovate web security. Knowing the cyber-criminal and his behavior before being the victim of an attack allows proactive measures to be taken in the face of an imminent attack.

The majority of solutions are not effective against the different attacks already known; however, solutions present that employ the classification of behavioral patterns, leaving aside traditional techniques and technologies. These solutions have not yet been applied to real contexts to measure and test their effectiveness. DDoS attacks stopping is already implementing automatic learning; most attacks are related to the application layer. Therefore, emphasis should be placed on generating policies and secure programming techniques to incur errors and leave open gaps that can be exploited by cybercriminals.

A honeypot can have many advantages; however, if it is not configured in a real environment and without considering whether its usefulness is adequate for the purpose for which it is to be used. It becomes obsolete, becoming a potential threat to the network infrastructure. Likewise, using an all-in-one honeypot such as T-Pot can be beneficial, considering that it should be implemented in an environment as real as possible and with constant monitoring. Although T-Pot Honeypot is free, open-source software, it consumes many resources. Therefore, it is advisable to allocate high storage capacity, information processing capacity, and bandwidth for data transmission, which gives the solution a medium implementation cost. On the other hand, it requires a high level of knowledge for its implementation and administration.

For this reason, it is observed that the most frequent attacks on the infrastructure of the Escuela Superior Politécnica de Chimborazo are DoS, represented at 39%, followed by the attack of SQL injection at 21%, Cross-site scripting (XSS) at 12%. Likewise, with 10% Brute Force Attacks, concluding that these are four categories of attacks with the highest number of perpetration records to the system or were used in greater proportion by cybercriminals to cause damage. Similarly, the country with the highest number of attacks during the period of implementation of the server in the infrastructure of the ESPOCH was the USA with the highest peak level of attacks, followed by Indonesia, then Brazil, Ghana, and Turkey, while there are countries with records of attacks in smaller numbers, but not less important than the others.

Regarding the classification of cognitive patterns acquired through Honeypots and ML algorithms for their processing, it is a new field that provides valuable information to understand better how cyber attackers or hackers operate and develop more effective countermeasures. The findings of this research can be shared with the cybersecurity community to develop advanced models that can identify patterns in real-time and mitigate them immediately or, in turn, implement more effective security measures.

It is necessary to develop instruments (psychological tests) aimed at hackers to have better results in future research and, likewise, to delimit a study population or sample, thus avoiding the bias that may be generated by applying the measurement instrument in hacker forums or sites. Currently, there is no single personality profile that describes all hackers, common traits have been investigated, such as different motivations, which can be criminal,



political, personal or criminal situations, this tells us that we should not stereotype all hackers.

Also, it generates a process for the automatic cleaning and construction of data from the patterns left by hackers in servers or honeypots; in this way, the database will be increased, which will obtain more accurate results by applying ML algorithms for processing, classification, and predictions.

ML algorithms, such as Neural Networks using a sequential model and Random Forest using 150 predictors, fit adequately to the training and test data as presented in Figure. 18. In the first analysis, the input parameters Country, Tool, IP, and Time have been used to model the 7 types of attacks described. When evaluating the Neural Network with the training data, an average F1-score of 99% is obtained, where the lowest value obtained F1-score is 92% for the class 'THE MAN IN THE MIDDLE.' While with test data that are unknown to the Network, a 98% average F1-score is obtained, with the lowest value of 80% for the class 'DEFACEMENT.' Compared to Random Forest, the mean F1-score is 100% with training data, while with test data, it is 98%. The lowest value obtained is 80% for the 'DEFACEMENT' class. The densities for these predictions are presented in Figure. 11. Comparing the performance of the 2 algorithms for the training data, Random Forest outperforms the Neural Network by 2%, while for the test data, the performance is similar.

For the second analysis, the input parameters Personality, Trait, Motivation, Country, Tool, and Time were used to model the 7 types of attacks described based on the personality characteristics associated with the hacker. Since, in the previous case, Random Forest performed better, this algorithm was used with 250 predictors for the model. The average F1-score obtained is 100% for the training data, while 93% is obtained for the test data. The lowest value obtained is 50% for the 'DEFACEMENT' class. As can be noticed in the 2 analyses, the 'DEFACEMENT' class presents a high complexity at the time of classification. A larger amount of data associated with this type of attack could be obtained to improve this result.

## 6.2. Future work

In the present investigation, patterns of external attacks were obtained worldwide. In future studies, the proposed scheme could be implemented to obtain information on the behavior of internal cybercriminals, that is, to know the cyber criminals of the same country, so that the security solution of a web server would be considered both externally and internally. Also, as future research, it is recommended to process the data through Machine Learning algorithms (Artificial Intelligence, AI) to determine patterns of behavior of cyber- criminals, determine an algorithm for training the data, and, based on the results, propose complementary security measures to those currently used for proactive, robust, efficient, and effective security.

Finally, the behavioral patterns of a cybercriminal can be very useful in web security to identify possible threats and prevent attacks. For this purpose, behavioral profiling is recommended to detect anomalous user behavior, identify vulnerabilities based on these patterns, and develop countermeasures to prevent or mitigate any possible intrusion into the web server.

As future research, the database elaborated from the data collected with the T-Pot Honey-pot should be available for the application of Machine Learning techniques to identify negative behavior patterns, develop models for early detection of threats, improve the response to potentially dangerous attacks, develop tools for the total security of a web server, among others.

It is recommended in future work to develop a model that converges all the tools applied in this research at a given time to have real-time results; in turn, it optimizes time, technological, and economic resources to have a complete and economical solution, unlike current defense systems that while it is true that already apply principles of artificial intelligence; however, they continue to work with known attack signatures leaving aside the entity responsible for the attacks as is the hacker or cybercriminal and with high costs of acquisition, maintenance and updating, unaffordable values for small and medium enterprises.

## REFERENCES

- [1] K. Y. P. S. G. Yoo and J. Kim, “Confidential information protection system for mobile devices,” *security and communication networks*, *Nature*, vol. 5, no. 12, pp. 1452–1461, 2012, doi: 10.1002/sec.516.
- [2] D.-J. Kim, K.-W. Chung, and K.-S. Hong, “Person authentication using face, teeth and voice modalities for mobile device security”, *Consumer Electronics, IEEE Transactions on*, vol. 56, no. 44, pp. 2678–2685, 2010. DOI: 10.1109/TCE.2010.5681156.
- [3] S. G. Yoo, S.-H. Kang, and J. Kim, “Sera: A secure energy and reliability aware data gathering for sensor networks”, vol. 73, May 2010, pp. 1–11. DOI: 10.1109/ICISA.2010.5480347.
- [4] S. Corporation, “The internet security threat report (istr)”, vol. 22, Apr. 2017, pp. 1–11.
- [5] C. Wang, T. Miu, X. Luo, and J. Wang, “Skyshield: A sketch-based defense system against application layer ddos attacks”, *IEEE Transactions on Information Forensics and Security*, vol. 13, pp. 559–573, Oct. 2017. DOI: 10.1109/TIFS.2017.2758754.
- [6] A. G. G. Cybenko and P. Thompson, *Cognitive hacking. Advances in Computers*, 2004.
- [7] A. K. Rajaram, B. C. Babu, *et al.*, “Api based security solutions for communication among web services”, in *2013 Fifth International Conference on Advanced Computing (ICoAC)*, IEEE, 2013, pp. 71–575.
- [8] M. Bugliesi, S. Calzavara, and R. Focardi, “Formal methods for web security”, *Journal of Logical and Algebraic Methods in Programming*, vol. 87, pp. 110–126, 2017.
- [9] R. Greenstadt and J. Beal, “Cognitive security for personal devices”, in *Proceedings of the 1st ACM workshop on Workshop on AI Sec*, 2008, pp. 27–30.
- [10] IBM, “Cognitive security”, 2016, p. 11.
- [11] M. R. Ogiela and L. Ogiela, “Security of cognitive information systems”, in *Information Technology Convergence: Security, Robotics, Automations and Communication*, Springer, 2013, pp. 427–433. DOI: 10.1007/978-94-007-6996-0\_44.
- [12] H. Armstrong and P. Forde, “Internet anonymity practices in computer crime”, *Information Management Computer Security*, vol. 11, pp. 209–215, Dec. 2003. DOI: 10.1108/09685220310500117.

- [13] S. Thill, “Considerations for a neuroscience-inspired approach to the design of artificial intelligent systems”, in *Artificial General Intelligence*, J. Schmidhuber, K. R. Thórisson, and M. Looks, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 247–254, ISBN: 978-3-642-22887-2.
- [14] S. Bailly, G. Meyfroidt, and J.-F. Timsit, “What’s new in icu in 2050: Big data and machine learning”, *Intensive Care Medicine*, vol. 44, Dec. 2017. DOI: 10.1007/s00134-017-5034-3.
- [15] M. Abadi, A. Agarwal, P. Barham, *et al.*, “Tensorflow: Large-scale machine learning on heterogeneous distributed systems”, *arXiv preprint arXiv:1603.04467*, 2016.
- [16] A. Holzinger, “Interactive machine learning for health informatics: When do we need the human-in-the-loop?”, *Brain Informatics*, vol. 3, no. 2, pp. 119–131, 2016.
- [17] A. L. Blum and P. Langley, “Selection of relevant features and examples in machine learning”, *Artificial intelligence*, vol. 97, no. 1-2, pp. 245–271, 1997.
- [18] J. Qiu, Q. Wu, G. Ding, Y. Xu, and S. Feng, “A survey of machine learning for big data processing”, *EURASIP Journal on Advances in Signal Processing*, vol. 2016, pp. 1–16, 2016.
- [19] D. O. H. Kwama Leonard Ogweno Obare Erick Oteyo, “Honey pot intrusion detection system”, *International Journal of Engineering Inventions*, vol. 4, pp. 28–41, 2014.
- [20] D. Bharti, “Honeypot-based intrusion detection system: A performance analysis”, Mar. 2016. DOI: 10.13140/RG.2.1.4599.9768.
- [21] N. Eliot, D. Kendall, and M. Brockway, “A flexible laboratory environment supporting honeypot deployment for teaching real-world cybersecurity skills”, *IEEE Access*, vol. 6, pp. 34 884–34 895, 2018.
- [22] X. Jiang, D. Xu, and Y.-M. Wang, “Collapsar: A vm-based honeyfarm and reverse honeyfarm architecture for network attack capture and detention”, *Journal of parallel and distributed computing*, vol. 66, no. 9, pp. 1165–1180, 2006.
- [23] V. Nicomette, M. Kaâniche, E. Alata, and M. Herrb, “Set-up and deployment of a high-interaction honeypot: Experiment and lessons learned”, *Journal in computer virology*, vol. 7, pp. 143–157, 2011.
- [24] I. Khan, H. Durad, and M. Alam, “Data analytics layer for high-interaction honeypots”, in *2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, IEEE, 2019, pp. 681–686.

- [25] S. Keele *et al.*, *Guidelines for performing systematic literature reviews in software engineering*, 2007.
- [26] V. R. Mouli and K. Jevitha, "Web services attacks and security-a systematic literature review", *Procedia Computer Science*, vol. 93, pp. 870–877, 2016.
- [27] M. Cai, Y. Lin, and W.-J. Zhang, "Study of the optimal number of rating bars in the likert scale", in *Proceedings of the 18th international conference on information integration and web-based applications and services*, 2016, pp. 193–198.
- [28] M. H. Kamarudin, C. Maple, T. Watson, and N. S. Safa, "A logitboost-based algorithm for detecting known and unknown web attacks", *IEEE Access*, vol. 5, pp. 26 190–26 200, 2017.
- [29] M. H. Kamarudin, C. Maple, T. Watson, N. S. Safa, *et al.*, "A new unified intrusion anomaly detection in identifying unseen web attacks", *Security and Communication Networks*, vol. 2017, 2017.
- [30] X. D. Hoang, "A website defacement detection method based on machine learning techniques", in *Proceedings of the 9th International Symposium on Information and Communication Technology*, 2018, pp. 443–448.
- [31] Y. Zhou and P. Wang, "An ensemble learning approach for xss attack detection with domain knowledge and threat intelligence", *Computers & Security*, vol. 82, pp. 261–269, 2019.
- [32] K. Ross, M. Moh, T.-S. Moh, and J. Yao, "Multi-source data analysis and evaluation of machine learning techniques for sql injection detection", in *Proceedings of the ACMSE 2018 Conference*, 2018, pp. 1–8.
- [33] D. Kar, S. Panigrahi, and S. Sundararajan, "Sqligot: Detecting sql injection attacks using graph of tokens and svm", *Computers & Security*, vol. 60, pp. 206–225, 2016.
- [34] D. Beckett, S. Sezer, and J. McCanny, "New sensing technique for detecting application layer ddos attacks targeting back-end database resources", in *2017 IEEE International Conference on Communications (ICC)*, IEEE, 2017, pp. 1–7.
- [35] H. Zhang, S. Wei, L. Ge, *et al.*, "Towards an integrated defense system for cyber security situation awareness experiment", in *Sensors and Systems for Space Applications VIII*, SPIE, vol. 9469, 2015, pp. 56–63.

- [36] M. Alsaleh, A. Alarifi, A. Alqahtani, and A. Al-Salman, "Visualizing web server attacks: Patterns in phpids logs", *Security and Communication Networks*, vol. 8, no. 11, pp. 1991–2003, 2015.
- [37] J. Shenbagam and P. Salini, "Vulnerability ontology for web applications to predict and classify attacks", in *2014 International Conference on Electronics, Communication and Computational Engineering (ICECCE)*, IEEE, 2014, pp. 268–272.
- [38] D. Hedin and A. Sabelfeld, "Web application security using jsflow", in *2015 17th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*, IEEE, 2015, pp. 16–19.
- [39] V. Clincy and H. Shahriar, "Web service injection attack detection", in *2017 12th International Conference for Internet Technology and Secured Transactions (ICITST)*, IEEE, 2017, pp. 173–178.
- [40] A. Z. M. Saleh, N. A. Rozali, A. G. Buja, K. A. Jalil, F. H. M. Ali, and T. F. A. Rahman, "A method for web application vulnerabilities detection by using boyer-moore string matching algorithm", *Procedia Computer Science*, vol. 72, pp. 112–121, 2015.
- [41] M. Srokosz, D. Rusinek, and B. Ksiezopolski, "A new waf-based architecture for protecting web applications against csrf attacks in malicious environment", in *2018 Federated Conference on Computer Science and Information Systems (FedCSIS)*, IEEE, 2018, pp. 391–395.
- [42] D. Fraunholz, D. Reti, S. Duque Anton, and H. D. Schotten, "Cloxy: A context-aware deception-as-a-service reverse proxy for web services", in *Proceedings of the 5th ACM workshop on moving target defense*, 2018, pp. 40–47.
- [43] T. Ishikawa and K. Sakurai, "Parameter manipulation attack prevention and detection by using web application deception proxy", in *Proceedings of the 11th international conference on ubiquitous information management and communication*, 2017, pp. 1–9.
- [44] A. M. Osman, A. Dafa-Allah, and A. A. M. Elhag, "Proposed security model for web based applications and services", in *2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE)*, IEEE, 2017, pp. 1–6.
- [45] A. Ekelhart, E. Kiesling, and K. Kurniawan, "Taming the logs-vocabularies for semantic security analysis", *Procedia Computer Science*, vol. 137, pp. 109–119, 2018.

- [46] S. Souissi, "Toward a novel classification-based attack detection and response architecture", in *2015 6th International Conference on the Network of the Future (NOF)*, IEEE, 2015, pp. 1–3.
- [47] H. Shahriar and W. Bond, "Towards an attack signature generation framework for intrusion detection systems", in *2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech)*, IEEE, 2017, pp. 597–603.
- [48] G. R. K. Rao, R. S. Prasad, and M. Ramesh, "Neutralizing cross-site scripting attacks using open source technologies", in *Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies*, 2016, pp. 1–6.
- [49] A. C. Perera, K. Kesavan, S. V. Bannakkotuwa, C. Liyanapathirana, and L. Rupasinghe, "E-commerce (web) application security: Defense against reconnaissance", in *2016 IEEE International Conference on Computer and Information Technology (CIT)*, IEEE, 2016, pp. 732–742.
- [50] A. W. Marashdih and Z. F. Zaaba, "Cross site scripting: Removing approaches in web application", *Procedia Computer Science*, vol. 124, pp. 647–655, 2017.
- [51] S. Gupta and B. Gupta, "Automated discovery of javascript code injection attacks in php web applications", *Procedia Computer Science*, vol. 78, pp. 82–87, 2016.
- [52] I. Parameshwaran, E. Budiando, S. Shinde, H. Dang, A. Sadhu, and P. Saxena, "Dexters: Robust testing platform for dom-based xss vulnerabilities", in *Proceedings of the 2015 10th Joint Meeting on Foundations of Software Engineering*, 2015, pp. 946–949.
- [53] K. S. Rao, N. Jain, N. Limaje, A. Gupta, M. Jain, and B. Menezes, "Two for the price of one: A combined browser defense against xss and clickjacking", in *2016 International Conference on Computing, Networking and Communications (ICNC)*, IEEE, 2016, pp. 1–6.
- [54] A. Niakanlahiji and J. H. Jafarian, "Webmtd: Defeating web code injection attacks using web element attribute mutation", in *Proceedings of the 2017 Workshop on Moving Target Defense*, 2017, pp. 17–26.

- [55] M. Taguinod, A. Doupé, Z. Zhao, and G.-J. Ahn, "Toward a moving target defense for web applications", in *2015 IEEE international conference on information reuse and integration*, IEEE, 2015, pp. 510–517.
- [56] B. Singh, K. Kumar, and A. Bhandari, "Simulation study of application layer ddos attack", in *2015 International Conference on Green Computing and Internet of Things (ICGCIoT)*, IEEE, 2015, pp. 893–898.
- [57] G. Deepa and P. S. Thilagam, "Securing web applications from injection and logic vulnerabilities: Approaches and challenges", *Information and Software Technology*, vol. 74, pp. 160–180, 2016.
- [58] S. Liang, Y. Zhang, B. Li, X. Guo, C. Jia, and Z. Liu, "Secureweb: Protecting sensitive information through the web browser extension with a security token", *Tsinghua science and Technology*, vol. 23, no. 5, pp. 526–538, 2018.
- [59] G. P. Bherde and M. Pund, "Recent attack prevention techniques in web service applications", in *2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT)*, IEEE, 2016, pp. 1174–1180.
- [60] A. Bartoli, A. De Lorenzo, E. Medvet, M. Faraguna, F. Tarlao, *et al.*, "A security-oriented analysis of web inclusions in the italian public administration", *Cybernetics and Information Technologies*, vol. 18, no. 4, pp. 94–110, 2018.
- [61] S. Souissi, "A novel response-oriented attack classification", in *2015 International Conference on Protocol Engineering (ICPE) and International Conference on New Technologies of Distributed Systems (NTDS)*, IEEE, 2015, pp. 1–6.
- [62] D. Kshirsagar, S. Sawant, A. Rathod, and S. Wathore, "Cpu load analysis & minimization for tcp syn flood detection", *Procedia Computer Science*, vol. 85, pp. 626–633, 2016.
- [63] O. Voitovych, O. Yuvkovetskyi, and L. Kupershtein, "Sql injection prevention system", in *2016 International Conference Radio Electronics & Info Communications (Ukr-MiCo)*, IEEE, 2016, pp. 1–4.
- [64] H.-C. Huang, Z.-K. Zhang, H.-W. Cheng, and S. W. Shieh, "Web application security: Threats, countermeasures, and pitfalls", *Computer*, vol. 50, no. 6, pp. 81–85, 2017.
- [65] H. Asghar, Z. Anwar, and K. Latif, "A deliberately insecure rdf-based semantic web application framework for teaching sparql/sparul injection attacks and defense mechanisms", *computers & security*, vol. 58, pp. 63–82, 2016.



- [66] M. Masango, F. Mouton, P. Antony, and B. Mangoale, "Web defacement and intrusion monitoring tool: Wdimt", in *2017 International Conference on Cyberworlds (CW)*, IEEE, 2017, pp. 72–79.
- [67] Y. Dong, Y. Zhang, H. Ma, *et al.*, "An adaptive system for detecting malicious queries in web attacks", *Science China Information Sciences*, vol. 61, pp. 1–16, 2018.
- [68] R. Hofstede, M. Jonker, A. Sperotto, and A. Pras, "Flow-based web application brute-force attack and compromise detection", *Journal of network and systems management*, vol. 25, pp. 735–758, 2017.
- [69] X. Han, N. Kheir, and D. Balzarotti, "Evaluation of deception-based web attacks detection", in *Proceedings of the 2017 Workshop on Moving Target Defense*, 2017, pp. 65–73.
- [70] A. Sultana and M. Jabbar, "Intelligent network intrusion detection system using data mining techniques", in *2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, IEEE, 2016, pp. 329–333.
- [71] K. Hong, Y. Kim, H. Choi, and J. Park, "Sdn-assisted slow http ddos attack defense method", *IEEE Communications Letters*, vol. 22, no. 4, pp. 688–691, 2017.
- [72] S. Laskar and D. Mishra, "Qualified vector match and merge algorithm (qvmma) for ddos prevention and mitigation", *Procedia Computer Science*, vol. 79, pp. 41–52, 2016.
- [73] C. Wang, T. T. Miu, X. Luo, and J. Wang, "Skyshield: A sketch-based defense system against application layer ddos attacks", *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 3, pp. 559–573, 2017.
- [74] H. Shan, Q. Wang, and C. Pu, "Tail attacks on web applications", in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017, pp. 1725–1739.
- [75] A. Agrawal, M. Alenezi, R. Kumar, and P. R. Khan, "Measuring the sustainable-security of web applications through a fuzzy-based integrated approach of ahp and topsis", *IEEE Access*, vol. 7, pp. 153 936–153 951, Dec. 2019. DOI: 10.1109/ACCESS.2019.2946776.

- [76] A. W. Marashdih, Z. F. Zaaba, K. Suwais, and N. A. Mohd, "Web application security: An investigation on static analysis with other algorithms to detect cross site scripting", *Procedia Computer Science*, vol. 161, pp. 1173–1181, 2019, The Fifth Information Systems International Conference, 23-24 July 2019, Surabaya, Indonesia, ISSN: 1877-0509. DOI: <https://doi.org/10.1016/j.procs.2019.11.230>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1877050919319416>.
- [77] R. Oliveira, M. Raga, N. Laranjeiro, and M. Vieira, "An approach for benchmarking the security of web service frameworks", *Future Generation Computer Systems*, vol. 110, Nov. 2019. DOI: [10.1016/j.future.2019.10.027](https://doi.org/10.1016/j.future.2019.10.027).
- [78] F. Tudela, J.-R. Higuera, J. Bermejo, J. A. Montalvo, and M. Argyros, "On combining static, dynamic and interactive analysis security testing tools to improve owasp top ten security vulnerability detection in web applications", *Applied Sciences*, vol. 10, Dec. 2020. DOI: [10.3390/app10249119](https://doi.org/10.3390/app10249119).
- [79] A. Tekerek, "A novel architecture for web-based attack detection using convolution neural network", *Computers Security*, vol. 100, Oct. 2020. DOI: [10.1016/j.cose.2020.102096](https://doi.org/10.1016/j.cose.2020.102096).
- [80] N. Kanagasabai, R. Subha, and A. Haldorai, "A novel method to detect and prevent sqlia using ontology to cloud web security", *Wireless Personal Communications*, vol. 117, Apr. 2021. DOI: [10.1007/s11277-020-07243-z](https://doi.org/10.1007/s11277-020-07243-z).
- [81] S. Sengan, D. David, M. Anam, *et al.*, "Cloud security service for identifying unauthorized user behaviour", *Cmc -Tech Science Press-*, vol. 70, pp. 2581–2600, Sep. 2021. DOI: [10.32604/cmc.2022.020213](https://doi.org/10.32604/cmc.2022.020213).
- [82] A. Shaheed, M. H. D. B. Kurdy, and D. Megias, "Web application firewall using machine learning and features engineering", *Sec. and Commun. Netw.*, vol. 2022, Jan. 2022, ISSN: 1939-0114. DOI: [10.1155/2022/5280158](https://doi.org/10.1155/2022/5280158). [Online]. Available: <https://doi.org/10.1155/2022/5280158>.
- [83] K. Barić, Đ. Vukić, and D. Jelaš, "Automatic conversion of server reports into client reports using sql server reporting services (ssrs) technology", vol. 1, pp. 301–306, Dec. 2023. DOI: [10.46793/ICEMIT23.301B](https://doi.org/10.46793/ICEMIT23.301B). [Online]. Available: <https://icemit.vpsblace.edu.rs/index.php/icemit/article/view/35>.

- [84] S. A. Wright, A. Sathyagiri, and R. Tayal, "Machine learning and the secure access service edge", in *2023 Congress in Computer Science, Computer Engineering, amp; Applied Computing (CSCE)*, Los Alamitos, CA, USA: IEEE Computer Society, Jul. 2023, pp. 2251–2258.
- [85] S. Venkatramulu, M. S. Waseem, A. Taneem, S. Y. Thoutam, S. Apuri, *et al.*, "Research on sql injection attacks using word embedding techniques and machine learning", *Journal of Sensors, IoT & Health Sciences*, vol. 2, no. 01, pp. 55–66, 2024.
- [86] T. d. J. Mateo Sanguino, "Enhancing security in industrial application development: Case study on self-generating artificial intelligence tools", *Applied Sciences*, vol. 14, no. 9, p. 3780, 2024.
- [87] M. Masango, F. Mouton, P. Antony, and B. Mangoale, "An approach for detecting web defacement with self-healing capabilities", *Transactions on Computational Science XXXII: Special Issue on Cybersecurity and Biometrics*, pp. 29–42, 2018.
- [88] S. Gupta and B. B. Gupta, "Js-san: Defense mechanism for html5-based web applications against javascript code injection vulnerabilities", *Security and Communication Networks*, vol. 9, no. 11, pp. 1477–1495, 2016.
- [89] B. Hou, K. Qian, L. Li, Y. Shi, L. Tao, and J. Liu, "Mongodb nosql injection analysis and detection", in *2016 IEEE 3rd International Conference on Cyber Security and Cloud Computing (CSCloud)*, IEEE, 2016, pp. 75–78.
- [90] L. Qian, Z. Zhu, J. Hu, and S. Liu, "Research of sql injection attack and prevention technology", in *2015 International Conference on Estimation, Detection and Information Fusion (ICEDIF)*, IEEE, 2015, pp. 303–306.
- [91] N. Patel and N. Shekokar, "Implementation of pattern matching algorithm to defend sqlia", *Procedia Computer Science*, vol. 45, pp. 453–459, 2015.
- [92] Y. Zhao, W. Zhang, Y. Feng, and B. Yu, "A classification detection algorithm based on joint entropy vector against application-layer ddos attack", *Security and Communication Networks*, vol. 2018, 2018.
- [93] R. Rojas, A. Muedas, and D. Mauricio, "Security maturity model of web applications for cyber attacks", in *Proceedings of the 3rd International Conference on Cryptography, Security and Privacy*, 2019, pp. 130–137.

- [94] M. Sachdeva, K. Kumar, and G. Singh, "A comprehensive approach to discriminate ddos attacks from flash events", *Journal of information security and applications*, vol. 26, pp. 8–22, 2016.
- [95] B. Mewara, S. Bairwa, J. Gajrani, and V. Jain, "Enhanced browser defense for reflected cross-site scripting", in *Proceedings of 3rd International Conference on Reliability, Infocom Technologies and Optimization*, IEEE, 2014, pp. 1–6.
- [96] D. K. R. Sekar, V. Gayathri, G. Anisha, K. S. Ravichandran, and R. MANIKANDAN, "Dynamic honeypot configuration for intrusion detection", May 2018, pp. 1397–1401. DOI: 10.1109/ICOEI.2018.8553956.
- [97] R. E. Mushtakov, D. S. Silnov, O. V. Tarakanov, and V. A. Bukharov, "Investigation of modern attacks using proxy honeypot", in *2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, Moscow: IEEE, Jan. 2018.
- [98] N. Innab, E. Alomairy, and L. Alsheddi, "Hybrid system between anomaly based detection system and honeypot to detect zero day attack", in *2018 21st Saudi Computer Society National Computer Conference (NCC)*, Riyadh: IEEE, Apr. 2018.
- [99] R. M. Arifianto, P. Sukarno, and E. M. Jadied, "An SSH honeypot architecture using port knocking and intrusion detection system", in *2018 6th International Conference on Information and Communication Technology (ICoICT)*, Bandung: IEEE, May 2018.
- [100] N. Naik, P. Jenkins, R. Cooke, and L. Yang, "Honeypots that bite back: A fuzzy technique for identifying and inhibiting fingerprinting attacks on low interaction honeypots", in *2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, Rio de Janeiro: IEEE, Jul. 2018.
- [101] O. Surnin, F. Hussain, R. Hussain, *et al.*, "Probabilistic estimation of honeypot detection in internet of things environment", in *2019 International Conference on Computing, Networking and Communications (ICNC)*, Honolulu, HI, USA: IEEE, Feb. 2019.
- [102] S. Kumar, B. Janet, and R. Eswari, "Multi platform honeypot for generation of cyber threat intelligence", in *2019 IEEE 9th International Conference on Advanced Computing (IACC)*, Tiruchirappalli, India: IEEE, Dec. 2019.
- [103] R. Vishwakarma and A. K. Jain, "A honeypot with machine learning based detection framework for defending IoT based botnet DDoS attacks", in *2019 3rd Interna-*

- tional Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India: IEEE, Apr. 2019.
- [104] D. Pliatsios, P. Sarigiannidis, T. Liatifis, K. Rompolos, and I. Siniosoglou, “A novel and interactive industrial control system honeypot for critical smart grid infrastructure”, in *2019 IEEE 24th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, Limassol, Cyprus: IEEE, Sep. 2019.
- [105] W. Fan, Z. Du, M. Smith-Creasey, and D. Fernandez, “HoneyDOC: An efficient honeypot architecture enabling all-round design”, *IEEE J. Sel. Areas Commun.*, vol. 37, no. 3, pp. 683–697, Mar. 2019.
- [106] E. Chovancova and N. Adam, “The security of heterogeneous systems based on cluster high-interaction hybrid honeypot”, in *2019 IEEE 23rd International Conference on Intelligent Engineering Systems (INES)*, Gödöllő, Hungary: IEEE, Apr. 2019.
- [107] C. Huang, J. Han, X. Zhang, and J. Liu, “Automatic identification of honeypot server using machine learning techniques”, *Secur. Commun. Netw.*, vol. 2019, pp. 1–8, Sep. 2019.
- [108] M. Dodson, A. R. Beresford, and M. Vingaard, “Using global honeypot networks to detect targeted ICS attacks”, in *2020 12th International Conference on Cyber Conflict (CyCon)*, Estonia: IEEE, May 2020.
- [109] M. M. Yamin, B. Katt, K. Sattar, and M. B. Ahmad, “Implementation of insider threat detection system using honeypot based sensors and threat analytics”, in *Lecture Notes in Networks and Systems*, ser. Lecture notes in networks and systems, Cham: Springer International Publishing, 2020, pp. 801–829.
- [110] P. Sharma and B. Nagpal, “HONEYDOS: A hybrid approach using data mining and honeypot to counter denial of service attack and malicious packets”, *Int. J. Inf. Technol.*, vol. 14, no. 2, pp. 837–846, Mar. 2022.
- [111] W. Zhang, B. Zhang, Z. Ying, H. He, and Z. Ding, “An iot honeynet based on multiport honeypots for capturing iot attacks”, *IEEE Internet of Things Journal*, vol. PP, pp. 1–1, Nov. 2019. DOI: 10.1109/JIOT.2019.2956173.
- [112] I. Muhamad Malik Matin and B. Rahardjo, “The use of honeypot in machine learning based on malware detection: A review”, Dec. 2020. DOI: 10.1109/CITSM50537.2020.9268794.

- [113] P. J. Chuang and T. C. Hung, “Enhanced attack blocking in iot environments: Engaging honeypots and machine learning in SDN OpenFlow switches”, *Journal of Applied Science and Engineering*, vol. 23, no. 1, pp. 163–173, 2020.
- [114] H. Taşçi, S. Gönen, M. A. Barişkan, G. Karacayilmaz, B. Alhan, and E. N. Yilmaz, “Password attack analysis over honeypot using machine learning password attack analysis”, *Turkish Journal of Mathematics and Computer Science*, vol. 13, no. 2, pp. 388–402, Dec. 2021.
- [115] F. D. Setiawan Sumadi, C. S. Kusuma Aditya, A. A. Maulana, S. Syaifuddin, and V. Suryani, “Semi-supervised approach for detecting distributed denial of service in SD-honeypot network environment”, *IAES Int. J. Artif. Intell. (IJ-AI)*, vol. 11, no. 3, p. 1094, Sep. 2022.
- [116] S. Prabakaran, R. Ramar, I. Hussain, *et al.*, “Predicting attack pattern via machine learning by exploiting stateful firewall as virtual network function in an SDN network”, *Sensors (Basel)*, vol. 22, no. 3, p. 709, Jan. 2022.
- [117] P. S. Aranjó, S. Maurya, C. Thakur, and M. Raju, “Threat prediction using honeypot and machine learning”, *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 10, no. 3, pp. 1838–1851, Mar. 2022.
- [118] A. K. Suleiman, A. Kayed, R. A. Shamat, V. Jagni, I. Obaid, and A. Awad, “A passive OS-Fingerprinting framework using honeypot”, in *2022 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSYS)*, Manama, Bahrain: IEEE, Jun. 2022.
- [119] A. Abdou, R. Sheatsley, Y. Beugin, T. Shipp, and P. McDaniel, “HoneyModels: Machine learning honeypots”, in *MILCOM 2021 - 2021 IEEE Military Communications Conference (MILCOM)*, San Diego, CA, USA: IEEE, Nov. 2021.
- [120] A. H. Khan, W. U. Khan, I. Hamid, A. W. Abbas, M. H. Chaudhry, and N. U. Arfeen, “Analysis and implementation of honeypot framework for enhancing network security”, *organization*, vol. 11, p. 12,
- [121] Z. Yao, H. Chen, K. Wang, H. Dong, J. Wan, and L. Zhang, “Design of lan security defense system based on honeypot technology”, in *Proceedings of the 2nd International Academic Conference on Blockchain, Information Technology and Smart Finance (ICBIS 2023)*, Atlantis Press, 2023, pp. 904–911.

- [122] T. Shivaprasad, A. Moulya, and N. Guruprasad, “Enhancing network security through a multi-layered honeypot architecture with integrated network monitoring tools”, in *2024 11th International Conference on Computing for Sustainable Global Development (INDIACom)*, IEEE, 2024, pp. 473–477.
- [123] M. Maranco *et al.*, “Enhanced honeypot security for intrusion detection and prevention systems using blockchain”, *World Journal of Advanced Research and Reviews*, vol. 22, no. 1, pp. 751–758, 2024.
- [124] I. M. M. Matin and B. Rahardjo, “The use of honeypot in machine learning based on malware detection: A review”, in *2020 8th International Conference on Cyber and IT Service Management (CITSM)*, IEEE, 2020, pp. 1–6.
- [125] F. D. S. Sumadi, C. S. K. Aditya, A. A. Maulana, S. Syaifuddin, and V. Suryani, “Semi-supervised approach for detecting distributed denial of service in sd-honeypot network environment”, *IAES International Journal of Artificial Intelligence*, vol. 11, no. 3, p. 1094, 2022.
- [126] W. Zhang, B. Zhang, Y. Zhou, H. He, and Z. Ding, “An iot honeynet based on multiport honeypots for capturing iot attacks”, *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 3991–3999, 2019.
- [127] D. Wu, C. Jennings, J. Terpenney, R. X. Gao, and S. Kumara, “A comparative study on machine learning algorithms for smart manufacturing: Tool wear prediction using random forests”, *Journal of Manufacturing Science and Engineering*, vol. 139, no. 7, p. 071 018, 2017.
- [128] R. Chiesa, S. Ducci, and S. Ciappi, *Profiling hackers: the science of criminal profiling as applied to the world of hacking*. CRC Press, 2008, vol. 49.
- [129] T. C. Summers, K. J. Lyytinen, T. Lingham, and E. A. Pierce, “How hackers think: A study of cybersecurity experts and their mental models”, *SSRN Electron. J.*, 2013.
- [130] M. Bachmann, “The risk propensity and rationality of computer hackers”, *International Journal of Cyber Criminology*, vol. 4, no. 2, pp. 643–656, 2010.
- [131] Q. Campbell and D. Kennedy, “The psychology of computer criminals”, in Sep. 2015, pp. 12.1–12.33, ISBN: 9781118134108. DOI: 10.1002/9781118851678.ch12.
- [132] A. Matulesy and N. H. Humaira, “Hacker personality profiles reviewed in terms of the big five personality traits”, *Psychology and Behavioral Sciences*, vol. 5, no. 6, pp. 137–142,

- [133] K. C. Seigfried-Spellar, N. Villacís-Vukadinović, and D. R. Lynam, "Computer criminal behavior is related to psychopathy and other antisocial behavior", *J. Crim. Justice*, vol. 51, pp. 67–73, Jul. 2017.
- [134] M. M. Al-Ajlouni, "Human resources and their tendency to information security crimes based on holland theory", *ng, Inf. Resour. Manag. J.*, vol. 31, no. 4, pp. 44–58, Oct. 2018.
- [135] S. Back, J. LaPrade, L. Shehadeh, and M. Kim, "Youth hackers and adult hackers in south korea: An application of cybercriminal profiling", in *2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, Stockholm, Sweden: IEEE, Jun. 2019.
- [136] J. Gaia, B. Ramamurthy, G. Sanders, *et al.*, "Psychological profiling of hacking potential", Jan. 2020. DOI: 10.24251/HICSS.2020.273.
- [137] K. Kioskli and N. Polemi, "Measuring psychosocial and behavioural factors improves attack potential estimates", in *2020 15th International Conference for Internet Technology and Secured Transactions (ICITST)*, London, United Kingdom: IEEE, Dec. 2020.
- [138] C. Stachl, *Predicting personality from patterns of behavior collected with smartphones*.
- [139] M. Bada and J. R. C. Nurse, "Profiling the cybercriminal: A systematic review of research", in *2021 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA)*, Dublin, Ireland: IEEE, Jun. 2021.
- [140] D. N. Jones, E. Padilla, S. R. Curtis, and C. Kiekintveld, "Network discovery and scanning strategies and the dark triad", *Comput. Human Behav.*, vol. 122, no. 106799, p. 106 799, Sep. 2021.
- [141] J. Mcalaney, "Are you anonymous? social-psychological processes of hacking groups", in *Cybersecurity and Cognitive Science*, Elsevier, 2022, pp. 139–155.
- [142] K. Yasuhara, D. Walnycky, I. Baggili, and A. Alhishwan, "The amorphous nature of hackers: An exploratory study", in *Annual ADFSL Conference on Digital Forensics, Security and Law*, 2022.



- [143] M. Weulen Kranenbarg, J.-L. van Gelder, A. J. Barends, and R. E. de Vries, “Is there a cybercriminal personality? comparing cyber offenders and offline offenders on HEXACO personality domains and their underlying facets”, *Comput. Human Behav.*, vol. 140, no. 107576, p. 107 576, Mar. 2023.
- [144] J. Gaia, B. Ramamurthy, G. Sanders, *et al.*, “Psychological profiling of hacking potential”, 2020.
- [145] R. Gabrys, A. Venkatesh, D. Silva, *et al.*, “Emotional state classification and related behaviors among cyber attackers”, in *56th Hawaii International Conference on System Sciences, HICSS 2023, Maui, Hawaii, USA, January 3-6, 2023*, T. X. Bui, Ed., ScholarSpace, 2023, pp. 846–855. [Online]. Available: <https://hdl.handle.net/10125/102735>.
- [146] C. Martinez-Santander, S. G. Yoo, and H. Moreno Aviles, “Analysis of traditional web security solutions and proposal of a web attacks cognitive patterns classifier architecture”, in *4th International Conference, CITI 2018*. Jan. 2018, pp. 186–198, ISBN: 978-3-030-00939-7. DOI: 10.1007/978-3-030-00940-3\_14.
- [147] A. A. Akinyelu, A. O. Adewumi, *et al.*, “Classification of phishing email using random forest machine learning technique”, *Journal of Applied Mathematics*, vol. 2014, 2014.
- [148] C. Martinez-Santander, H. Oswaldo, and M. Hernández, “Analysis of intrusions into computer systems using honeypots”, *International Journal of Intelligent Systems and Applications in Engineering*, vol. 11, pp. 461–472, Dec. 2023.
- [149] “Copyright 2022 by the institute of electrical and electronics engineers, inc. all rights reserved”, in *2022 20th International Conference on Optical Communications and Networks (ICOON)*, Shenzhen, China: IEEE, Aug. 2022.

## **ANNEXES**

### **Annexe A**

The evolution from Traditional to Intelligent Web Security: Systematic Literature Review

### **Annexe B**

Analysis of Traditional Web Security Solutions and Proposal of a Web Attacks Cognitive Patterns Classifier Architecture

### **Annexe C**

Analysis of Intrusions into Computer Systems using Honeypots

### **Annexe D**

Classification of cognitive patterns of hackers using Machine Learning

### **Annexe E**

Hacker personality test

### **Annexe F**

Tabulation of the Hacker personality test

### **Annexe G**

BFQ personality manual

### **Annexe H**

T-score using the scale

