International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Analysis of Intrusions into Computer Systems using Honeypots

Carlos José Martínez S [ab]*, Hugo Oswaldo Moreno A [c] and Myriam Beatriz Hernández A [a]

Submitted: 11/02/2023 **Revised:** 14/04/2023 **Accepted:** 13/05/2023

Summary: Among the different tools and techniques to detect, monitor and analyze the behavior of cybercriminals in a controlled environment is the use of Honeypots. This computer security technique consists of creating a trap or decoy that pretends to be a valuable target for cybercriminals, such as, for example, important data, applications or web servers, among others. Once the attacker enters this environment, it records all the activities and behaviors performed by them, this in turn, allows the detection and proactive response to future attacks. Likewise, they allow to reduce false positives and identify vulnerabilities of the systems. On the other hand, the implementation of these tools implies a potential risk, because they can be used by attackers to obtain information about the defenses of the institution or organization and, even more, if it is not installed correctly, leaving the system and sensitive data exposed. In this sense, there are many high and low interaction Honeypots, such as, for example, Honeyd, Dioneda, Capture-HPC, KFSensor, INetSim. Cowrie, Honeytrap, Amun, Glastopf, Contop, among others. Finally, T-Pot Honeypot is considered as a set of computers (All in one) containing both high and low interaction Honeypots, with a unique ability to customize and manage honeypots, that is, it is a powerful and effective tool in the fight against attackers and computer security threats. In addition, it allows the obtaining of behavioral data of cybercriminals to later determine patterns that allow improving Web Security.

Keywords: *Honeypot, TPot, Web Security, Attacks, Cyber-criminals*

1. Introduction

Currently, cyberattacks have increased, [1]–[3]making information and communication security a critical issue due to the constant increase in cyber threats and sophistication of attacks, malware, denial of service (DDoS), among others. 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 to the point of finding new ways to exploit vulnerabilities present in web servers or overcome existing security tools and technologies. Therefore, web security remains a major challenge and of vital attention to keep us updated and be aware of new threats as well as the proposed solutions for the case. [4]–[7][8][9]

The International Telecommunication Union (ITU), according to the fourth edition of the cybersecurity index says that, Ecuador is ranked 119 out of 182 countries with compromised information security, that is, it represents a percentage of 26.3% of failures or vulnerabilities, the most common attacks carried out are those of ransomware-type malware, where, the Telecommunications Regulation and Control Agency (ARCOTEL) qualifies as a medium level of security risks. For this reason, the Ministry of Telecommunications and the Information Society[10][11] has incorporated a national cybersecurity strategy, where the government set the guidelines for security in cyberspace. Whose purpose is to ensure that the information of all Ecuadorian users is protected and are not subject to cyberattacks. Among the main strategies is the execution of six points: "1 governance and national coordination, 2. cyber resilience, 3. prevention and fight against cybercrime, 4. national cyber defense, 5. cybersecurity skills and capabilities, 6. international cooperation".[12]

Complemented by the above, there are several ways to protect information and ensure security in both servers and data infrastructures through the use of honeypots that are useful tools to detect vulnerabilities and strengthen network security [13], [14] . In addition, honeypots have the ability to repel attacks against them and distract the attacker with a decoy network similar to the original, that [15]way the cyber-criminal is confused and believes that he is entering the real system and that he is gaining environments controlled by the owner, without imagining that he is inside a trap. This environment is prepared with all the information of the resources, tools, techniques and vulnerabilities that the attacker assumes that the network has and will use its resources for their possible attacks. Thehoneypot will take advantage to collect all that

[a] *Facultad de Ingeniería de Sistemas, Escuela Politécnica Nacional, Quito, Ecuador.*
[b] *Carrera de Medicina, Universidad Católica de Cuenca, Cuenca, Ecuador.*
[c] *Facultad de Informática y Electrónica, Escuela Superior Politécnica de Chimborazo, Riobamba, Ecuador.*
carlos.martinez03@epn.edu.ec, h_moreno@espoch.edu.ec, myriam.hernandez@epn.edu.ec

information for in the future to implement a secure network that can prevent, detect and respond to thesepossible threats.

In summary, honeypots have many applications and uses, because another of their functions is the diversion of suspicious traffic from critical systems, receiving warnings that will delay or stop attacks that may cause damage to the real infrastructure of the company or organization. According to the above, these technologies or tools represent multiple advantages and applications in the field of web security given their ability to collect valuable information about the tactics and tools used by attackers in a system[16]. This can help identify patterns and trends in the behavior of cybercriminals, which is useful for the field of investigation combined with computer security, cognitive security or threat intelligence and digital forensics. In general, the use of honeypots can provide valuable and unique information that cannot be obtained through other research methods, so the objective of this article is to review the state of the art on the subject addressed with respect to honeypots and the implementation of T-P[17], [18][19]–[21]ot Honeypot to demonstrate its effectiveness, Analyze and present the main results on the behavior of cybercriminals.

Background

Honeypots

There are honeypots [22]-[23] high and low interaction, the main difference is that high interaction honeypots are more complex and expensive to implement, however, they provide a holistic and real view of the techniques and patterns used by attackers. In contrast, low-interaction honeypots are easier and cheaper, which translates to less detailed insight and can be easily detected by attackers[24][25]. The selection for the application of either a high or low interaction honeypot will depend on the needs[26], objective and resources of the organization or the research project. Table 1. shows a comparison between these two technologies.

Table 1. High and low interaction honeypot matching

Low Interaction Honeypot	High Interaction Honeypot
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 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.

Source: Author of the research, (Martínez C.2023)

T-Pot Honeypot is a hive of honeypots, the main difference between other solutions of this type is the ability to emulate multiple services and operating systems, integration with SIEM (Security Information and Event Management), installer and complete documentation. Below is a comparative table of some honeypots used in different research projects based on the literature consulted:

Table 2. Ddescription of honeypot types

Honeypot	Interaction	Customization	Integration	Virtualization	Difficulty of use	Integration with SIEM
T-Pot Honeypot	High / Low	Loud	Yes	Yes	Media	Yes
Honeyd	Loud	Loud	No	Yes	Loud	No
Dioneda	Loud	Media	Yes	Yes	Media	No

KFSensor	Loud	Loud	Yes	No	Loud	Yes
Glastopf	Casualty	Media	No	Yes	Media	No
Amun	Casualty	Loud	Yes	Yes	Loud	No

Source: Author of the research, (Martínez C.2023)

Machine Learning and its application in Honeypots

Honeypot within security is a very important resource, it was designed to be attacked and examine its possible attacks, however, cybercriminals always put a step ahead in what has to do with security, they look for ways and strategies to boycott the security that is provided to the systems. According to him, he [14], proposes to update honeypot techniques so that they are not recognized by hackers and silently capture their data to counter it. For the purpose, the researcher proposed to use intelligent techniques such as Machine Learning and thus automatically check whether the honeypot is being executed or not by the server, likewise, this proposal applies the random forest algorithm with three characteristics: from the network application layer and the system where the data from known public systems were obtained and the model was used to measure its efficiency, resulting in a high value of 0.93 i.e. the algorithm is suitable for this purpose.

On the other hand, [27] it raises in its research work the detection of malware using honeypot, as malware is known is one of the information security risks on the rise. Millions of malicious programs are detected daily, giving rise to a greater number of threats, these are usually found in Trojan malware and adware, consider that security devices such as firewalls, IDS and antivirus do not detect malware. On the other hand, in this research alternatives have been found to detect malware using Machine Learning with honeypot, where honeypot serves as a trap for suspicious packages and machine learning detects malware, in the end the proposal is to integrate the Decision Tree algorithm and Support Vector Machine (SVM) as classification algorithms for malware detection.

As in any computer system, cloud networks are not out of reach of attacks, since they are exposed and exploited with greater emphasis by data pirates, because the cloud covers a greater number of data and it is large companies, houses of studies or organizations. In addition, this kind of network has no restrictions for the user, that is, he has freedom of use. In this regard, [2] he proposed to use a Honeynet system consisting of three durable honeypots to detect attacks and to determine the recorded behavior of the Honeynet, we experimented with three trained machine learning methods, the three models using are Naive Bayes, SVM and Random Forest to more accurately classify the new incoming data from these trap networks as malicious.

Alhan [28] , expresses that with digitalization in the world people have moved away from their social lives and they are dependent on technology, this dependence has led to cyberthreats and therefore attacks, for this reason, Alhan B proposes in this research work the detection of attacks in real time using HoneyPi and Machine Learning. In this environment what was done is a comprehensive security architecture that integrates HoneyPi in a Raspberry Pi based on open software and hardware with low costs, for the evaluation of the data base will be used two Machine Learning algorithms first uses Naive Bayes the reason for using this algorithm is for its high degree of accuracy in few data and does not work by learning but does it directly by That reason is very inexpensive. Next, the LSTM machine learning algorithm is used, which provides continuous and sequential learning on the Datasets it processes.

In the research, it was proposed to use an industrial honeypot to detect botnet attackers [29]by deliberately creating resources on the network, to solve the problem of closely monitoring and capturing the behavior of botnet attacks. As a result, the content reviewed and written to log files, these records are quickly classified with great accuracy using machine learning operations. Obtaining an increase in productivity and improving the stability of smart factories. This study proposes a botnet detection model that combines honeypots and machine learning designed specifically for smart factories. The study was conducted with hardware that simulates a smart factory environment.

According [30], to the compromised SSH servers used by cybercriminals are usually difficult to detect, however, by analyzing system logs, organizations can learn about them and thus improve performance in their security, this would be the case of companies with several SSH servers. Manual log analysis can be tedious and high-speed networks require better mechanisms to detect system anomalies. In this post, a compromised SSH session performing malicious activities is discovered, for this, they used a flow-based approach and machine learning techniques to detect compromised sessions from the perspective of the technique is transmission, where individual packets are not analyzed; Therefore, it works best on high-speed networks. The data comes from distributed honeypots. The paper also describes machine learning methods with appropriate parameter and function selection techniques, including showing a real-time detection model tested on a public server.

In the research, the use of[31] a T-Pot Honeypot is proposed, in order to discover what types of attacks and malicious intentions are carried out by cybercriminals and different strategies used to take control of the network. In the present research, the authors expose the configuration and results obtained by the honeypot and subsequently the application of a machine learning algorithm to predict the type of threat. The prioritized system services were: Apache Webserver, MYSQL, FTP and SMTP, one of the limitations determined is that a good hacker will be able to determine that he is attacking a honeypot. Likewise, GMM clustering algorithms were used to group the IP addresses found and subsequently label them as malicious and non-malicious.

In recent years, with the sharp increase in DDoS attacks by botnets on IoT, the security of these devices has become one of the most worrying issues in network security. Many security approaches have been proposed in this area, but they are not yet able to deal with the new variant of IoT malware, i.e. zero-day attack. This research presents a honeypot-based approach that uses learning techniques to detect malware. Data generated for efficient and dynamic training of a machine learning model. In this sense, it can be considered as an effective start in the fight against zero-day DDoS attacks, which is now an open challenge for IoT protection.[32]

The ELK Stack

For [33] ELK it is a combination of open source (Elasticsearch, Logstash and Kibana) that runs in a virtual environment. It has a comprehensive approach to data analysis and consolidation, each component fulfills its role Elasticsearch is a search engine based on Apache, while Logstash collects necessary records and sends to Elasticsearch where it is responsible for converting it into JSO format, finally, Kibana is responsible for visualizing this data in tables, graphs or maps. For all these advantages, ELK was used to build a great security log analysis system for companies, with low costs in installing commercial products and helping that companies that are starting up do not strive to create other systems of records that take longer and are less effective. ELK has proven to have a record logging time and provides various types of visualization tool helping security administrators, therefore, the ELK stack proves to be a powerful element in computer security analysis.

In the research[34], the authors propose the integration of the ELK, SIEM and PACK stack as an alternative security solution since joining the three or more components would obtain a robustness in the security of companies that use large amounts of data. This integration allows you to use each of the characteristics of the components such as: security, machine learning and alerts. This whole proposal focuses on the protection of private and openly processed data on certain types of platforms, so it forces stakeholders to pay close attention to data security and look for protection alternatives.

In [35], the authors refer to the operation of ELK combined for efficient log analysis and provide user-friendly tools. The log systems that are supported by ELK are made to analyze large amounts of data and also facilitate the monitoring of processes and their calculation by means of an interactive interface. Being an open source has many facilities for the analysis of records. Elasticsearch is used as an indexing, storage, and retrieval mechanism. The Log acts as input slider and dicer and output writer and finally Kibana performs the visualization through dashboard. For this article, the implementation of ELK was efficiently geo-known to web users for registration.

Chen et al. [36] exposes that the technology presents problems in the collection of Docker cluster records, such as low efficiency, weak application and poor stability, the proposal is to adopt the ELK, Filebeaty Kafka for the design of a system collector and analyzer of Docker containers records supported with ELK, to perform a rapid deployment to perform the collection of records in real time, The filtering and sending of data visualization and analysis has greatly improved the efficiency of the work of the staff. Resulting in good real-time performance, stability and high availability

2. Methodology

For the execution of this project, an infrastructure was implemented in the dedicated for the purpose that consisted of a first firewall owned by the service provider CEDIA that together with ESPOCH managed the creation of a tunnel with the aim of protecting the data infrastructure and services of the Higher Education Institution. In the server used for the purpose was an HP ProLiant DL360 Gen9, it was located in the Faculty of Computer Science and Electronics FIE in the ESPOCH, the installation process of the T-Pot was developed, which previously downloaded the image. GitHub ISO for post-installation.

Afterwards, all the services that ESPOCH such as website, databases, among others, were replicated. The use of the different open ports Table **3** on a web server was also guaranteed , in such a way, incentivize the attacker to perform different attacks such as DDoS, XSS, SQL injection, brute force, CSRF and buffer overflow attacks.

Table 3. Description of protocols

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

Source: Author of the research, (Martínez C.2023)

The server addresses were posted on forums and sites of cyber-criminals or hackers[1]. The system has been online for more than six months the information is collected from the 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.

3. Results and Discussion

The main results obtained from the present study are discussed below.

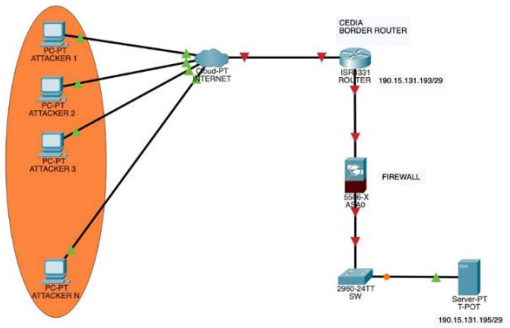

Fig 1. Server infrastructure

Source: Author of the research, (Martínez C.2023)

Fig 2. Login

Source: Author of the research, (Martínez C.2023)

The administration of the T-Pot Honeypot was done from SSH and HTTPS connection through the server address, figure 2 shows the main screen of the server, in addition, through the ELK Stack stack the graphic administration as the visualization of the results becomes super simple, however, you must have a high level of knowledge for additional or personal configurations according to the purpose of the investigation. On the other hand, figure 3 shows the server activity with the active ports of the different services mounted on the Honeypots hive.

Fig 3. Initialized ports for honeypot

Source: Author of the research, (Martínez C.2023)

Among the services mounted on the web server, the main website of the ESPOCH was replicated, to meet this objective the Httrack tool was used, first the information from the website was downloaded, in this process it took approximately 36 hours. Subsequently, we proceeded to mount on the web server previously mounted on the physical server implemented, this can be seen in figure 4.

Fig 4. IP of the ESPOCH home page

Source: Author of the research, (Martínez C.2023)

Among the main results after the time more or less than six months of the server deployed in the ESPOCH infrastructure, it was observed that, the attacks perpetrated D DoS occupy the first place with a large number of attacks of 39%. Followed by the SQLInjection attack 21%, Cross-site scripting (XSS) 12%. Likewise, with 10% Brute Force Attacks, attacks preferred by cyber-criminals as observed in a previous work [4], all this is seen in figure 5.

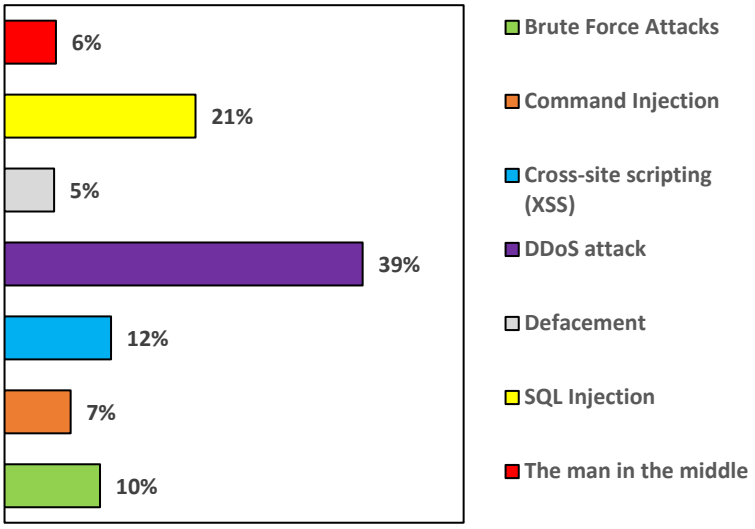

Fig 5. Attacks on the server deployed in the ESPOCH infrastructure

Source: Author of the research, (Martínez C.2023)

Also, an important fact that can lead to valuable conclusions of the present study is the time zone figure. 6, the T-Pot Honeypot recorded 18 time zones as favorites among which in first place is the GMT-7 zone with 26.6% belonging to Asian countries such as Indonesia Thailand, Cambodia, Singapore. Then follows GMT+7 representing the USA with 24.8%. Followed with 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 Islamizes, Syria and GMT-3 Brazil. It should be considered that most cybercriminals use VPNs, however, it can also be a pattern that improves the security of a web server, since time zones are repeated, which allows to determine that they do not change VPNs constantly, rather they hide the source address, but they are not permanently randomizing the IPs used for attacks.

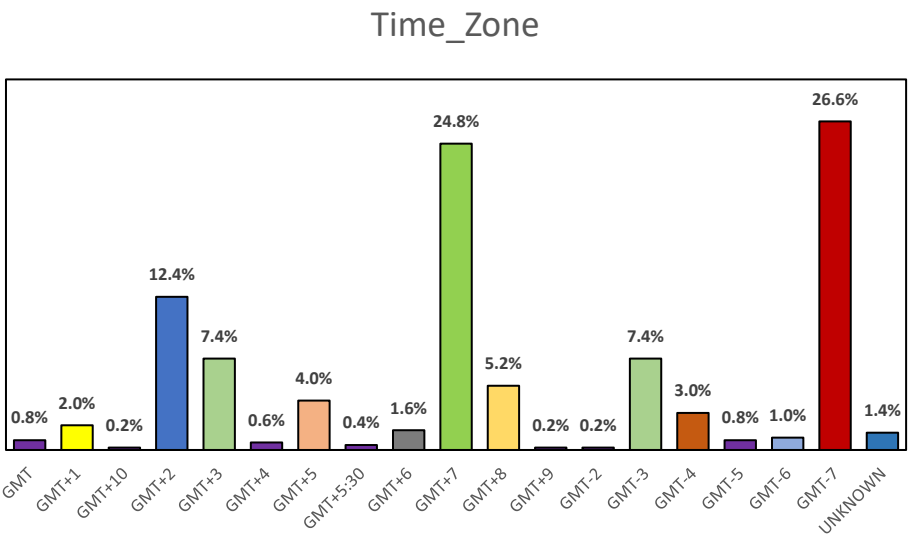

Fig 6. Time zone from which the Attacks were launched

Source: Author of the research, (Martínez C.2023)

The T-Pot allows you to configure a server farm in addition to the aforementioned web page, the results of figure 7 show that there is a preference for attacking web servers with Apache 46.20% of the attacks were against this server. Among the data recorded in the T-pot Honeypot is 32.20% to more than two servers which is registered with the tag of unknown. After with 12.60% this nginx, this last server has become popular in the last decade given its robustness.

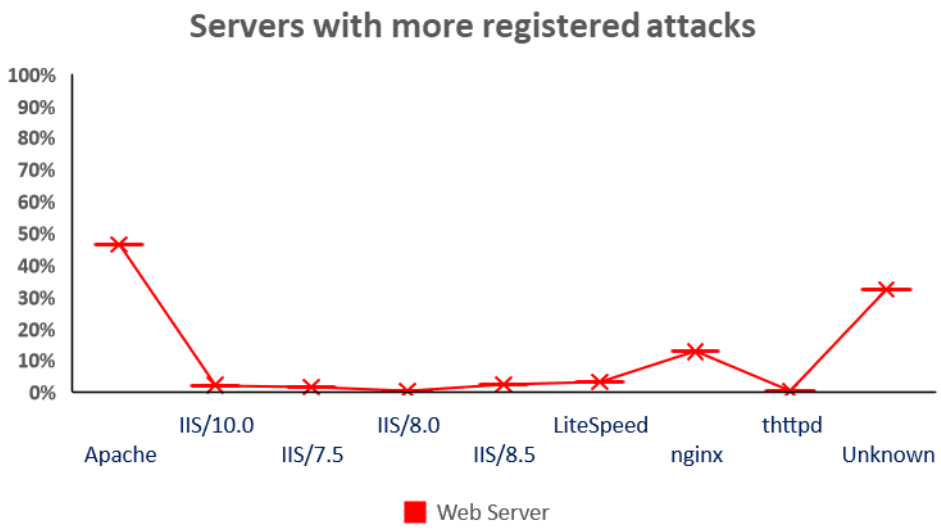

Fig 7 server with more registered attacks

Source: Author of the research, (Martínez C.2023)

As stated above, launching a computer attack is not limited by nationality or geographical location, therefore, it is difficult to determine with certainty which country carries out the largest number of computer attacks and even more so when there are tools to mask the source IP or the use of VPNs to change location virtually. However, some countries such as the United States, China, Russia, North Korea and Iran are constantly accused of espionage and carrying out multiple large-scale computer attacks. In addition, some developing countries may also be involved in illegal cyber activities, as seen in Figure 8, in the implemented T-Pot, it registers the United States as the country with the most attacks carried out, followed by Indonesia, Brazil and Iran.

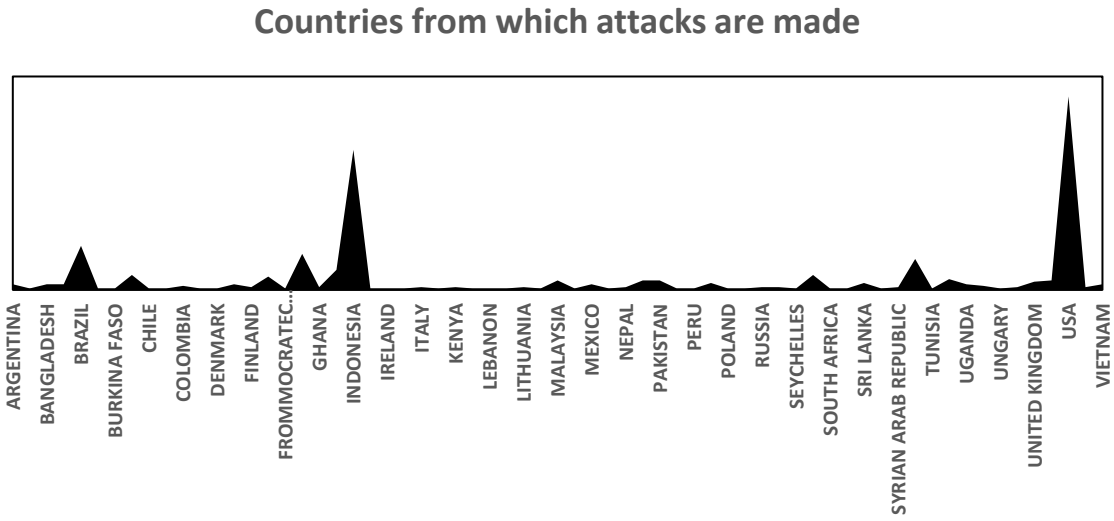

Fig 8 Countries from which attacks are made

Source: Author of the research, (Martínez C.2023)

In the research of Back et al. , states that, in the [37]fourth position of the motivation that cybercriminals have when they belong to a group or want to demonstrate their leadership in it, the ego is present. The T-Pot Honeypot allows you to register the footprint of cyber-criminals by signing (Nickname), most left their signature, this is shown in table number 5, important data to create more detailed profiles of computer criminals. There are many cybercriminal forums in which these nicknames are exposed to develop an online reputation, in web security this data could serve largely to associate with these profiles and generate security measures based on that information.

Tabla 5. Database of Nickname of Cybercriminals

Database of Nicknames of Cybercriminals

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	Xyp3r2667	Black_TeamX
SABUNMANDI CYBER TI	Mr.Kro0oz.305	finistro lammer	Crystal_MSF	AKINCILAR	x153	evil-net
Admeral zino_dz	Mr.Kro0oz.305	YIIX103	NULL SQU4D	KAKEGURAI	greecye	T-Freak
Hamza Anonime	Faisal 1337	aDriv4	Salim Alk	pi.hack	IDOLSEC Team	0xRiko
Black_X12	theMx0nday	finistro lammer	Akbar dravinky	Trenggalek Cyber Army	Unknown Al	./G1L4N6_ST86
Sofian X35 Dz	Mr.ToKeiChun69	Imam	DR1D1	r1dA	Bloos3rpent	Micin
LahBodoAmat	White Cyber Illusion	MiSh	PohonSagu	R0cket	Mister_XID	Mr.Rm19
FoursFROMath Team	Moroccan Revolution	TUNOVATO	Royal Battler BD	Yemen Cyber Army	chinafans	John lover
AlfabetoVirtual	0x1998	Clash Hackers	R13S	Simsimi	Zer0FaulT	Phenix-TN
NDA	KrdSec	iccarmy	AnoaGhost	LucasOwna	djebbaranon	Bla3k_D3vil
Collapse Gang	Family Attack Cyber	NDA	Tn Jones	RexShelby	Mr.Jenskins	Trio Ngefly
pr0s3x	Wedus	Ghost Hunter Illusion	RzkyO	Banjarnegara Xploit	dhuaa	/Dandi.alien
Fatal Error	CLAN_X12	TurkHackTeam	AnoaGhost	Hadii6666h0$T	Zodiacxs	Mrb3hz4d
m1kesecurity	ro0t-M8n	Umam1337	LahBodoAmat	Zyyy	Mr.XaaD	LISTEN
Black_X12	Paraná Cyber Mafia	D3D0T	Zeerx7	B3g0k[Kurdish Hacker]	FRK48	Dijehaji
KrdSec	D.R.S Dz Team	sh007	K4TSUY4-GH05T	V1ruz3L	VenoRyan	l0v3ly17
Arch1999	Panataran	./MrTahuSumedang	M15T4k3	Typical Idiot Security	Black-Python	MrGalang
Mr.Kro0oz.305	/Rayzky_	Toro	Trenggalek Cyber Army	F.Z MalaikatHati	ALHOSANE	ZoRRoKiN
mr.anFROMrson	Sc0rp10n.DZ	AnonSec Team	Salman Hacker	./s3nt1n3L	FurkHan	ONE HAT CYBER TEAM
K4PUYU4K	Ren4Sploit	z3ran gaza hack3er	./Cyber00t	Mamad Warning	Panataran	9 ECTUS_TM
Stoupid!	Fuck MALAYSIA	Indonesia Attacker #O	Jebe Was Here!	Fii	SH3LL_INS3RT3R	Adel SabrH
Mosawi--	Boky992	CyberFrost	Matigan1337	./Newbie4rt_ID	./Juba_Dz	Netr4LizeR
Babacang07	OTm4n3	Mr.Froggy	/KeyzNet	RootAyyıldız	Centra7x	Mr V
Melody-x48	Clash Hackers	luxe	KosameAmegai	/Synchronizer	ProxssHunter	Pul
Ahd404	SDM21	LucasOwna	Al Catraz	Illusion Silent Killer	bilat	xNot_RespondinGx
TheLevelSevenCrew	Cubjrnet7	TheZero	ZERA	AhmadBlocker	1975 Team	AnonyKs_xD
KURD ELECTRONIC TEA	Lumajang Xploit	27F HaCkEr	Frustrated Hackers BD	0xJoshua	Albania Attacker	Krypton-X
/unn0rmaL	Unknown Al FUCKTARD	Galang10	Tn Bekicot	Astra	Verad	AnonCoFROMrs
7870a	Shield iran	Dhen Bhocil	VaporEkJet	./Mr.xWanz403x	Hevnen	Mr#

Source: Author of the research, (Martínez C.2023)

4. Conclusions and Future Work

With the digital evolution of companies and public institutions nationwide, users become perfect targets for cybercriminals, the same ones who take advantage of the web to remain anonymous and carry out attacks. Institutions in order to minimize the likelihood of being victims have increased and improved the security system in their institutions. That is why, ESPOCH is not outside the security detection systems, since it allowed the installation of the server infrastructure for the detection of attacks and thus have knowledge and seek solutions to threats.

Currently, the security measures of a web server are provided by firewall rules or WAF devices intended for packet analysis. Analyzing data other than those mentioned will improve and innovate web security, knowing the cyber-criminal and their behavior before being victims of an attack, allows generating proactive measures against an eminent attack.

A honeypot can have many advantages, however, if it is not configured in a real environment and without taking into account if the usefulness of the same is adapted to the purpose for which it is going to be used, it is 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 must be implemented in an environment as real as possible and with constant supervision. The T-Pot Honeypot despite being free open source software, consumes a lot of resources. Therefore, it is advisable to allocate high storage capacity as well as information processing capacity and bandwidth for data transmission, this makes the solution have a medium implementation cost. On the other hand, it requires a high level of knowledge for its implementation and management.

For this reason it is observed that the most frequent attacks on the infrastructure of the Polytechnic School of Chimborazo is D DoS represented with 39%, followed by the attack of SQLInjection 21%, Cross-site scripting (XSS) 12%. Likewise, with 10% Brute Force Attacks, concluding that there are four categories of attacks with the highest amount of perpetration record to the system or were used in greater proportion by cybercriminals to cause damage. In the same way, the country with the highest number of attacks was obtained during the period of implementation of the server in the ESPOCH infrastructure, it was the USA with the highest peak at the level of attacks, followed by Indonesia, then by Brazil, Ghana and Turkey, at the same time that there are countries with records of attacks in smaller numbers, but no less important than the others.

In the present research patterns of external attacks were obtained worldwide, in future works the proposed scheme could be implemented to obtain information on the behavior of internal cybercriminals, that is, to know the

computer criminals of the same country, in this way the security solution of a web server would be contemplated both externally and internally. Also, as future work it is recommended to process the data through Machine Learning algorithms (Artificial Intelligence AI) to determine patterns of behavior of cybercriminals, determine an algorithm for training the data and from the results, propose security measures complementary to those currently used for proactive, robust security, efficient and effective.

Finally, the behavior patterns of a cybercriminal can be very useful in the field of web security to identify potential threats and prevent attacks. For this, it is recommended to develop behavioral profiles that will detect anomalous behavior of users, as well as identification of vulnerabilities from these patterns and develop countermeasures to prevent or mitigate any possible interference of the web server.

As future work it is recommended to have the database elaborated from the data collected with the T-Pot Honeypot so that through the application of Machine Learning techniques you can identify malicious behavior patterns, develop models of early detection of threats, improve the response to potentially dangerous attacks, develop tools for the total security of a web server, among others.

References

[1] C. Martínez Santander, S. G. Yoo, and H. O. Moreno, "Analysis of traditional web security solutions and proposal of a web attacks cognitive patterns classifier architecture," in Communications in Computer and Information Science, 2018, pp. 186–198. doi: 10.1007/978-3-030-00940-3_14.

[2] M. Marydas and J. N. Varshapriya, "A Cloud based Honeynet System for Attack Detection using Machine Learning Techniques," International Research Journal of Engineering and Technology, vol. 330, no. July, pp. 330–335, 2019.

[3] C. I. Rene and J. Abdullah, "Malicious Code Intrusion Detection using Machine Learning And Indicators of Compromise," Ijcsis, vol. 15, no. September, pp. 160–171, 2017, [Online]. Available: http://www.academia.edu/download/55691451/Journal_of_Computer_Science_IJCSIS_September_2017_Full_Volume.pdf#page=173

[4] J. Martinez, "The evolution from Traditional to Intelligent Web Security : Systematic Literature Review," 2020.

[5] R. G. T. Thilagam, "A Review on the Effectiveness of Machine Learning and Deep Learning Algorithms for Cyber Security," Archives of Computational Methods in Engineering, no. 0123456789, 2020, doi: 10.1007/s11831-020-09478-2.

[6] S. Mazumdar and J. Wang, Guide to Vulnerability Analysis for Computer Networks and Systems, no. September. 2018. doi: 10.1007/978-3-319-92624-7.

[7] I. H. Sarker, A. S. M. Kayes, S. Badsha, H. Alqahtani, P. Watters, and A. Ng, "Cybersecurity data science: an overview from machine learning perspective," J Big Data, vol. 7, no. 1, 2020, doi: 10.1186/s40537-020-00318-5.

[8] G. E. Rodríguez, J. G. Torres, P. Flores, and D. E. Benavides, "Cross-site scripting (XSS) attacks and mitigation: A survey," Computer Networks, vol. 166, p. 106960, 2020, doi: 10.1016/j.comnet.2019.106960.

[9] 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, no. 4, pp. 735–758, Oct. 2017, doi: 10.1007/s10922-017-9421-4.

[10] Itu, "Global Cybersecurity Index 2020 Measuring commitment to cybersecurity Acknowledgements."

[11] Telecommunications Regulation and Control Agency, "EC-2022-031_RANSOMWARE-ELBIE_V1," Quito, 2022.

[12] Ministry of Telecommunications and Information Society, "CIBERSEGURIDAD DEL ECUADOR," 2022. Accessed: Apr. 24, 2023. [Online]. Available: https://asobanca.org.ec/wp-content/uploads/2022/08/ESTRATEGIA-NACIONAL-DE-CIBERSEGURIDAD-DEL-ECUADOR-2022481.pdf

[13] Flores I. and Quintana J., "SYSTEM OF DETECTION OF COMPUTER ATTACKS TO ENTERPRISE DATA NETWORKS SUPPORTED IN HONEYPOTS," Repositorio Universidad de Cartagena, pp. 1–95, 2018.

[14] C. Huang, J. Han, X. Zhang, and J. Liu, "Automatic identification of honeypot server using machine learning techniques," Security and Communication Networks, vol. 2019, 2019, doi: 10.1155/2019/2627608.

[15] Dimitrios Pliatsios, Panagiotis Sarigiannidis, Thanasis Liati fis, Konstantinos Rompolos, and Ilias Siniosoglou, "A Novel and Interactive Industrial Control System Honeypot for Critical Smart Grid Infrastructure," in IEEE COMMUNICATIONS SOCIETY, INSTITUTR OF ELECRICAL AND ELECTRONICS ENGENEERS, 2019, pp. 1–6.

[16] Ridho Maulana Ari fianto, Parman Sukarno, and Erwid Musthofa Jadied, "An SSH Honeypot Architecture Using Port Knocking and Intrusion Detection System," in 2018 6th International

Conference on Information and Communication Technology (ICoICT), 2018, pp. 409–415.

[17] W. Ruiz Martínez, "Analysis of Machine Learning techniques applied to computer cybersecurity to improve the detection of intrusions and anomalous behaviors on the Web."

[18] L. Dirección et al. , "CUADERNOS DE LA GUARDIA CIVIL," REVISTA DE SEGURIDAD PÚBLICA 3a ÉPOCA EDITOR-IN-CHIEF EDITOR-IN-CHIEF EDITORIAL BOARD, [Online]. Available: http://publicacionesoficiales.boe.es/

[19] M. R. Ogiela and L. Ogiela, "Security of Cognitive Information Systems," Springer, 2013, pp. 427–433.

[20] R. Greenstadt and J. Beal, "Cognitive security for personal devices," ACM, 2008, pp. 27–30.

[21] W. Kinsner, "Towards cognitive security systems," in 2012 IEEE 11th International Conference on Cognitive Informatics and Cognitive Computing, IEEE, Aug. 2012, pp. 539–539. doi: 10.1109/ICCI-CC.2012.6311207.

[22] M. N. Hoda, I. Bharati Vidyapeeth's Institute of Computer Applications and Management (New Delhi, Institute of Electrical and Electronics Engineers. Delhi Section, and I. International Conference on Computing for Sustainable Global Development (3rd : 2016 : New Delhi, "Honeypot-Based Intrusion Detection System: A Performance Analysis," Honeypot-Based Intrusion Detection System: A Performance Analysis, vol. 16, no. 18, pp. 3947–3951, 2016.

[23] 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. 34884–34895, Jun. 2018, doi: 10.1109/ACCESS.2018.2850839.

[24] X. Jiang, D. Xu, and Y.-M. Wang, "Collapsar: A VM-Based Honeyfarm and Reverse Honeyfarm Architecture for Network Attack Capture and Detention."

[25] V. Nicomette et al. , "Set-up and deployment of a high-interaction honeypot: experiment and lessons learned Set-up and deployment of a high-interaction honeypot: experiment and lessons learned Set-up and deployment of a high-interaction honeypot: Experiment and lessons learned," Journal in Computer Virology, vol. 7, no. 2, 2011, doi: 10.1007/s11416-010-0144-2ï.

[26] Iqra Khan, Hanif Durad, and Masoom Alam, Data Analytics Layer For high-interaction Honeypots. 2019.

[27] I. M. M. Matin and B. Rahardjo, "Malware Detection Using Honeypot and Machine Learning," 2019 7th International Conference on Cyber and IT Service Management, CITSM 2019, 2019, doi: 10.1109/CITSM47753.2019.8965419.

[28] B. Alhan, S. Gönen, G. Karacayilmaz, M. A. Barişkan, and E. N. Yilmaz, "Real-Time Cyber Attack Detection Over HoneyPi Using Machine Learning," Tehnicki Vjesnik, vol. 29, no. 4, pp. 1394–1401, 2022, doi: 10.17559/TV-20210523121614.

[29] S. Lee, A. Abdullah, N. Z. Jhanjhi, and S. H. Kok, "Honeypot Coupled Machine Learning Model for Botnet Detection and Classification in IoT Smart Factory – An Investigation," MATEC Web of Conferences, vol. 335, p. 04003, 2021, doi: 10.1051/matecconf/202133504003.

[30] G. K. Sadasivam, C. Hota, and B. Anand, "Detection of Severe SSH Attacks Using Honeypot Servers and Machine Learning Techniques," Software Networking, vol. 2017, no. 1, pp. 79–100, 2017, doi: 10.13052/jsn2445-9739.2017.005.

[31] V. Mehta, P. Bahadur, M. Kapoor, P. Singh, and S. Rajpoot, "Threat prediction using honeypot and machine learning," 2015 1st International Conference on Futuristic Trends in Computational Analysis and Knowledge Management, ABLAZE 2015, pp. 278–282, 2015, doi: 10.1109/ABLAZE.2015.7155011.

[32] R. Vishwakarma and A. K. Jain, "A honeypot with machine learning based detection framework for defending IoT based botnet DDoS attacks," Proceedings of the International Conference on Trends in Electronics and Informatics, ICOEI 2019, no. Icoei, pp. 1019–1024, 2019, doi: 10.1109/ICOEI.2019.8862720.

[33] S. Son and Y. Kwon, "Performance of ELK Stack and Commercial System in Security Log Analysis," in IEEE 13th Malaysia International Conference on Communications (MICC), The Puteri pacific and Johor Bahru, Eds., Malaysia, 2017, pp. 187–190.

[34] F. Abdou, M. L. Salihi, and M. F. Nanne, "Toward a Secure ELK Stack." [Online]. Available: https://sites.google.com/site/ijcsis/

[35] T. Prakash, M. Kakkar, and K. Patel, "Geo-Identification of Web Users through Logs using ELK Stack," in 2016 6th International conference- cloud system and Big Data Engineering, IEEE, 2016, pp. 606–610.

[36] L. Chen, J. Liu, M. Xian, and H. Wang, "Docker container log collection and analysis system based on ELK," in Proceedings - 2020 International Conference on Computer Information and Big Data Applications, CIBDA 2020, Institute of Electrical and Electronics Engineers Inc., Apr. 2020, pp. 317–320. doi: 10.1109/CIBDA50819.2020.00078.

[37] S. Back, J. Laprade, L. Shehadeh, and M. Kim, "Youth hackers and adult hackers in south korea: An application of cybercriminal profiling," in Proceedings - 4th IEEE European Symposium on Security and Privacy Workshops, EUROS and PW 2019, Institute of Electrical and Electronics Engineers Inc., Jun. 2019, pp. 410–413. doi: 10.1109/EuroSPW.2019.00052.