Article

Not peer-reviewed version

Honeypots in Cybersecurity: Their Analysis, Evaluation and Importance

Zlatan Moric , Leo Mršić * , Zdravko Kunić , Goran Đambić

Posted Date: 13 August 2024

doi: 10.20944/preprints202408.0946.v1

Keywords: honeypot; cybersecurity; attack; threat; analysis

 doi:10.20944/preprints202408.0946.v1

Article

Honeypots in Cybersecurity: Their Analysis, Evaluation and Importance

Zlatan Morić [1], Leo Mršić [2,*], Zdravko Kunić [2] and Goran Đambić [3]

[1] Department of Cybersecurity and System Engineering, Algebra University, 10000 Zagreb, Croatia
[2] Department of Information Systems and Business Analytics, Algebra University, 10000 Zagreb, Croatia
[3] Department of Software Engineering, Algebra University, 10000 Zagreb, Croatia
* Correspondence: leo.mrsic@algebra.hr

Abstract: Honeypots are vital security tools that enable cybersecurity specialists to scrutinize the tactics employed by attackers. By doing so, they can get valuable insights into existing and upcoming internet dangers. Cyber-attacks are in a perpetual evolution, increasingly sophisticated and tailored to exploit specific vulnerabilities. Hence, it is imperative for both organizations and individuals not only to consider but also implement methods to safeguard themselves against cyberattacks and malware. This paper assesses a specific group of honeypot systems by considering detection range, emulation accuracy, data quality, dependability, scalability, performance, extensibility, installation options, setup difficulty, and maintenance requirements. Furthermore, it provides conclusions regarding the relative effectiveness of diverse honeypot systems in different situations. As a part of the research, it analyzes the latest progress and innovations in honeypot development, including enhancements in machine learning, heightened automation, integration with other security tools, cloud-based honeypots, and deception technology by using simulated assaults to showcase the logging and recording capabilities of these honeypots in capturing attack information. After carefully analyzing each honeypot, each has unique advantages, disadvantages, and special characteristics.

Keywords: honeypot; cybersecurity; attack; threat; analysis

1. Introduction

Honeypots have become essential in cybersecurity because they can analyze attacker methods and collect vital intelligence on existing and developing threats. These decoy systems aim to lure malicious actors, enabling cybersecurity professionals to analyze attack methods and strengthen defense mechanisms. This introduction explores the significance of honeypots in cybersecurity, utilizing recent research findings to emphasize their usefulness and development.

Cyber threats are constantly changing, becoming increasingly complex and focused. Given the continually evolving nature of these threats, it is imperative to implement sophisticated measures to safeguard sensitive information and systems. Honeypots serve as a proactive defense strategy by imitating vulnerable systems to attract attackers, thereby allowing the gathering of valuable data on attack patterns, techniques, and tools employed by cyber adversaries [1]. This data is crucial for the development of solid cybersecurity measures and the improvement of the overall security position of organizations.

An essential benefit of honeypots is their capacity to offer immediate and accurate observations of attackers' actions. For example, the implementation of Digital Twin honeypots has been demonstrated to improve threat intelligence by adjusting to different network scenarios and identifying sophisticated and persistent threats [2]. In addition, honeypots can detect and anticipate fingerprinting attacks in real time, a crucial feature preventing attackers from identifying and circumventing these deceptive systems [3]. These capabilities are necessary for preserving the integrity and efficiency of honeypots as a defensive tool.

 doi:10.20944/preprints202408.0946.v1

Furthermore, honeypots are paramount in educational environments as they significantly enhance students' understanding and knowledge of cybersecurity. Research has shown that engaging in honeypot activities dramatically improves students' comprehension of cybersecurity principles and technical abilities [4]. Through active involvement with honeypots, students acquire hands-on experience in identifying and examining cyber threats, a skill of immense value for their prospective careers as cybersecurity experts.

Recent research has also emphasized the incorporation of honeypots with other security technologies. For instance, integrating honeypots with machine learning algorithms and blockchain frameworks has created advanced systems that can accurately forecast and reduce cyber threats more efficiently [5]. These innovations showcase the continuous progress in honeypot technology and its capacity to transform cybersecurity practices.

Moreover, honeypots offer crucial observations into the tactics, techniques, and procedures (TTPs) employed by attackers. Researchers can collect extensive data on attacks by strategically placing honeypots in different environments, including phishing, scamming, and account hijacking [6]. This information is crucial for developing precise defense strategies and enhancing threat detection capabilities.

Recent research has highlighted the importance of strategically deploying honeypots in various geographical areas to maximize their efficacy. Researchers can enhance threat intelligence and develop more precise profiles of attackers by examining attack patterns from honeypots spread across different geographical locations [7]. This method is especially advantageous for developing early warning systems to detect advanced cyber threats.

Adaptive honeypots, which can modify their behavior based on attackers' actions, have also demonstrated significant efficacy. These systems employ methods such as reinforcement learning to actively involve attackers and gather valuable information while reducing the chances of being detected [8]. Adaptability ensures honeypots' continued relevance and effectiveness in a rapidly evolving threat environment.

Exploring honeypots in Internet of Things (IoT) environments has yielded promising outcomes. Scientists have created complex and diverse honeypot systems that become more advanced over time, depending on the behavior of the attackers they observe. These systems are precious for comprehending the distinct difficulties and dangers linked to IoT devices [9].

In addition, honeypots have been incorporated into software-defined networks (SDNs) to improve their implementation and administration. This methodology enables the consolidation of control and minimizes the occurrence of incorrect alerts, thereby enhancing the overall efficiency and efficacy of honeypot systems [10].

Honeypots have been employed in industrial cybersecurity to replicate industrial control systems (ICS) and identify cyberattacks. These systems can imitate different protocols and devices, allowing for a better understanding of attacker tactics and enhancing critical infrastructure security [11].

Honeypots are essential for detecting and analyzing malware. By capturing and analyzing malevolent activities, they aid researchers in comprehending and mitigating novel and developing threats. This feature is crucial for addressing zero-day attacks and other sophisticated cyber threats [12].

The ongoing advancements in honeypot technology are driving progress in cybersecurity. Researchers have utilized large language models (LLMs) to construct more authentic and dynamic honeypots. These systems can effectively involve human attackers, offering a more thorough understanding of their strategies and methods [13].

Researchers have also investigated the implementation of honeypots in wireless networks, which has yielded substantial advantages. Wi-Fi honeypots can promptly identify and react to unauthorized individuals, thereby improving the security of both residential and corporate networks [14].

Furthermore, game-theoretic methods have been utilized to enhance the effectiveness of honeypot placement and deployment strategies. Researchers can improve the effectiveness of honeypot systems and minimize risks by simulating the interactions between defenders and attackers [15].

 doi:10.20944/preprints202408.0946.v1

The combination of blockchain technology and honeypots has demonstrated potential. Utilizing blockchain-based frameworks can improve the security and dependability of honeypot systems by creating unchangeable records of identified threats and more robust defense strategies [16].

Virtual honeypots have been employed to create and simulate intrusion detection systems. These systems can analyze vast amounts of data and detect cybersecurity vulnerabilities, rendering them valuable assets for network security [17].

Researchers have utilized cognitive models to gain insight into the impact of various honeypot configurations on attacker behavior. Researchers can enhance the effectiveness of honeypot systems by analyzing the choices made by adversaries in different situations [18].

Academic and research institutions have recently implemented honeypots to collect cyberattack data and improve security education. These deployments offer valuable insights into attack patterns and contribute to the training of future cybersecurity professionals [19].

Ultimately, honeypots are essential instruments in the realm of cybersecurity. Due to their capacity to replicate actual systems, gather comprehensive attack data, and incorporate cutting-edge technologies, they have become a fundamental component of contemporary cybersecurity strategies. Continual research and development in honeypot technology will further improve their efficacy, guaranteeing their importance as a crucial element in combating cyber threats [20].

2. Research Goal and Methods

This research thoroughly examines the significance of honeypots in cybersecurity and compares specific solutions - it examines various categories of honeypots, outlining their benefits and drawbacks, and presents a comparative analysis of well-known solutions. The study was conducted on multiple aspects, including levels of interaction (low and high), supported operating systems, services, and attack detection and analysis capabilities. Since this research focuses on the implementation and comparison of specific honeypot solutions, as well as some changes to best practices and recommendations, the findings will help enhance the existing procedures in the field of cyber security. Through a comparative evaluation, analysis, and case study of the implementation of honeypot solutions, this research aims to uncover valuable information that organizations can utilize to mitigate risk and formulate innovative defense strategies.

We assess the significance of honeypots as a defensive measure, examine various types of honeypots, and evaluate existing solutions based on their effectiveness, implementation complexity, scalability, and maintenance demands. The research findings will enhance the comprehension and utilization of the honeypot technology for cybersecurity experts. The paper aims to assist organizations in selecting the optimal solution for their needs by conducting a thorough analysis. This analysis enables organizations to identify the most appropriate solution that aligns with their specific security protocol requirements. This paper will additionally present a comprehensive summary of various honeypot solutions, enabling organizations to discern the merits and drawbacks of each solution. Furthermore, this paper will offer practical recommendations for effectively implementing honeypots.

To prepare this paper, extensive research was undertaken to thoroughly comprehend the significance of honeypots in cybersecurity and compare the currently available solutions. Following this goal, the existing literature was examined, specifically books and articles that address the relevant topic. This paper utilized the following scientific research methods:

- The computer collection method is used to gather pertinent literature and web sources listed in the literature list, serving the purpose of writing this paper.
- The inductive method was employed in this study to analyze specific individual cases and gather new information about the role and significance of honeypots.
- This paper employed a deductive approach to analyzing the comprehensive aspects of honeypots in cybersecurity. This method involves drawing conclusions based on assumed truths and suggestions.

 doi:10.20944/preprints202408.0946.v1

- This paper employs analytical techniques to deconstruct intricate concepts, evaluations, and conclusions into their fundamental elements. Specifically, it applies these methods to analyze an individual honeypot solution according to predetermined criteria.
- The comparison method assesses the similarities and differences between the same or related facts. It allows for identifying standard features or distinctions when evaluating different honeypot solutions.

In the next section, we'll briefly overview honeypot types before evaluating honeypot solutions and comparing them.

3. Honeypot Types

Honeypots can offer dependable insights into the evolution of threats. They offer details regarding attack vectors, exploits, and malware. Hackers continuously refine their methods of unauthorized access, and the honeypot detects and identifies emerging risks and intrusions. Honeypots are also capable of detecting and capturing insider threats.

Various categories of honeypots can be employed to discern distinct categories of threats. The different interpretations of honeypots are contingent upon the specific nature of the danger being mitigated. Each of them plays a role in a comprehensive and efficient cybersecurity strategy. Email traps, also known as spam traps, involve the placement of a fictitious email address in a concealed location that can only be detected by an automated email address collector. Given that the address serves no purpose other than being a spam trap, it is unequivocal that any email received on it is spam. Messages with identical content to those sent to the spam trap can be automatically blocked, and the sender's original IP can be included in the ban list.

Malware honeypots share similarities with spam honeypots, but their primary focus is investigating malware that can potentially target an organization's systems. Spam honeypots are created to trap undesired bots and other forms of automated traffic.

Honeypots are highly efficient in identifying familiar and unfamiliar threats, enhancing threat detection capabilities. Honeypots can entice potential threats and surveil their activities by establishing an enticing environment for attackers. This can offer a valuable understanding of attackers' behavior and techniques to penetrate systems. Examining attacks is another crucial element of the honeypot. One can gain insight into their strategies and methods by carefully examining the behaviors exhibited by assailants within a regulated setting. Typical honeypot positioning can be viewed in Figure 1:

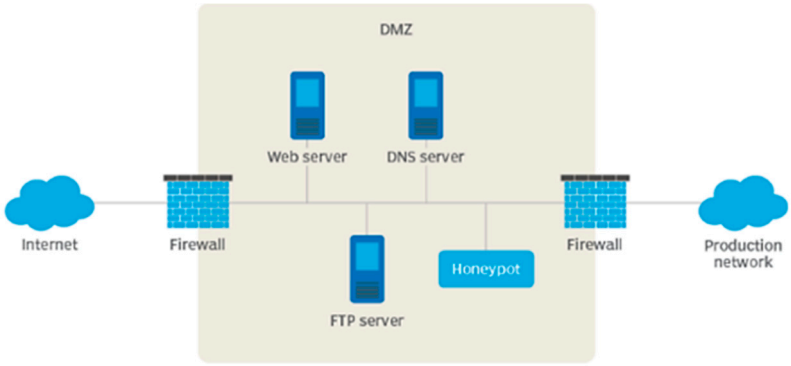

Figure 1. Honeypot network positioning [21].

Honeypots used for research are precious systems because they can collect data to analyze the methods employed in malicious activities. Setting up, maintaining, and collecting substantial data for research honeypots is complex. However, they are precious security tools that enhance forensic analysis capabilities. Furthermore, honeypots can enhance seizure prevention in addition to the insights gained from research.

Integrating production honeypots into current security systems enhances the capacity to detect and respond to attacks. Organizations utilize production honeypots to strengthen protection and

 doi:10.20944/preprints202408.0946.v1

mitigate risk levels. It enables you to improve the security of your system promptly. Due to its reduced functionality compared to a research honeypot, the development and implementation of this type of honeypot are typically less complex. However, it can detect various attack techniques, although it reveals less information about the perpetrator than a research honeypot.

Honeypots can be classified into various categories. Practically, honeypots are typically classified according to their capacity for interaction. Let's now describe the three main types by these criteria: high-interaction, low-interaction, and hybrid honeypots.

3.1. High-Interaction Honeypots

High-interaction honeypots aim to entice intruders by operating actual services. They are costlier and necessitate more frequent upkeep. While they possess certain benefits, they can also catalyze security vulnerabilities. Ensuring complete isolation and implementing all necessary safety precautions is imperative for networks with honeypots and high interactions. Alternatively, if an unauthorized individual gains access to a genuine system, they can infiltrate the Honeypot, potentially resulting in security vulnerabilities for the entire network.

3.2. Low-Interaction Honeypots

Honeypot and low-interaction techniques have a restricted level of interaction. These systems accurately replicate services and operating systems. The Honeypot system restricts intruder activities to the extent of emulation. Low-interaction honeypots are systems that imitate particular TCP/IP model protocols. Communication of information or signals from one point to another. Honeypots, specifically those with low interaction, do not store genuine or significant personal information. They require the minimum system specifications.

3.3. Hybrid Honeypots

A hybrid honeypot combines two with varying degrees of user engagement. The combination is a secure solution as it allows for the simultaneous use of both types of Honeypots, complementing each other and mitigating their respective limitations. An optimal approach would involve utilizing a Honeypot system with low interaction capabilities. It involves a significant level of interaction. A honeypot is a cybersecurity technique used to attract and deceive potential attackers, allowing organizations to gather information about their tactics and techniques.

The primary objective of employing hybrid honeypots is to leverage the scalability of low-interaction honeypots and the potential for high-interaction honeypot processing to gather comprehensive information on attack activities within extensive corporate networks. This enables successful analysis and evaluation of potential new attacks. Hybrid honeypot systems utilize low-interaction honeypots as access gateways for high-interaction honeypots and integrate various techniques and components. This allows them to target attack activities specifically, reducing analysis time and faster attack response.

4. Honeypot Solutions and Comparison Criteria

The honeypot solutions that we selected for this research are the most used, like:

- SSH honeypots – Kippo and Cowrite.
- HTTP honeypots – Glasstopf, Nodepot and Google Hack Honeypot.
- WordPress honeypots – Formidable Honeypot, Blackhole For Bad Bots, Wordpot.
- Database honeypots – ElasticHoney, HoneyMysql and MongoDB-HoneyProxy.
- Email honeypots – Honeymail, Mailoney, and SpamHAT.
- IoT honeypots – HoneyThing.
- Other types are Dionaea, Honeypot-ftp, HoneyNTP, Thug, and Canarytokens.

The selected honeypot solutions were compared based on the criteria used to evaluate them. The requirements are established using the data from [22] and the authors' experience. The requirements encompass the detection scope, emulation accuracy, quality of data collected, reliability, scalability

 doi:10.20944/preprints202408.0946.v1

and performance, extensibility, embeddability, complexity of setup and use, and maintenance requirements. The detection range refers to the extent of various attack vectors that a honeypot can identify. Emulation accuracy refers to the similarity between the emulated version of an application or service in a honeypot and its authentic counterpart. This criterion is irrelevant for high-interaction honeypots, which real applications or services provide. The data collected measures its quality within the context of the security system, as provided by the solution. Reliability pertains to the consistency and dependability of the solution when subjected to a heavy workload. Scalability and performance assess the capacity of a single instance of a tool and the capability to distribute the workload among multiple parallel processes or computing nodes within a single honeypot system. Extensibility is a metric that quantifies the difficulty level involved in expanding a tool's current capabilities to meet specific needs. Embeddability refers to the capacity of a solution to seamlessly integrate with other tools or operate within a more extensive system by utilizing an interface offered by a honeypot. Setup and usage complexity pertains to the intricacy of utilizing, overseeing, and configuring a solution.

5. Comparative Evaluation and Results Analysis

This section of the paper will present a comparative evaluation and analysis, which will be based on the criteria outlined in the previous section. As we can see in Table 1, there is a summary of the evaluation results in line with the specified criteria:

Honeyd is a versatile tool that can imitate any service by utilizing TCP or UDP protocols, as previously stated about the criteria for detection range. Dionae possesses a versatile detection range, and its protocols are integrated as modules. The target scope of Cowrie is particular. The detection scope of Amun is versatile, with distinct modules responsible for handling individual services and their vulnerabilities. Glasstopf simulates vulnerable web applications in its detection scope, whereas Kippo is limited to imitating the Secure Shell service. Thug can emulate a web browser and its plugins, explicitly identifying and mitigating threats propagated through malicious websites.

Both Honeyd and Dionaea exhibit high levels of emulation accuracy. Based on this criterion, there is a lack of specific data for Cowrie. However, considering its resemblance to Kippo, we can infer that it yields similar outcomes. Amun's emulation achieves high accuracy, with each service being emulated to the extent required to handle specific malware exploits and downloads and detect attacks on applications. Glasstopf emulation exhibits exceptional accuracy, enabling the attacker to execute typical website attacks and deliver the anticipated content to the target. Kippo and Thug both excel in terms of emulation accuracy.

Based on the quality criteria gathered, it is essential to highlight that Honeyd is poor quality. It only captures fundamental information such as timestamps, source and destination IPs, protocols, and source and destination ports. Dionaea possesses exceptional data collection capabilities, and its primary objective is to preserve the malware employed by the attacker. Cowrie logs all activities on the honeypot, including the attacker's commands and keystrokes, offering detailed information about the aggressor's actions. The data collected by Amun, Glasstopf, and Kippo is of high quality.

Regarding reliability, it is essential to note that Honeyd functions consistently without any reported issues. Dionaea is highly commendable according to this standard, displaying remarkable stability over an extended duration. Cowrie is compatible with and easily integrated into popular threat intelligence platforms and SIEM tools. Amun demonstrates exceptional reliability, exhibiting no stuttering or unexpected shutdowns during testing. According to this standard, Glasstopf and Kippo both meet the excellent criteria.

Regarding the complexity of installation and use for Honeyd, it is essential to note that the installation process requires basic technical knowledge but is not challenging. According to this standard for Dionaea, the installation may appear complex, but it is thoroughly documented and essentially involves replicating specific steps.

 doi:10.20944/preprints202408.0946.v1

Table 1. Results of evaluation based on defined criteria.

Title 1	Honeyd	Dionaea	Cowrie	Amun	Glasstopf	Kippo	Thug
Detection range	Multifunctional	Multifunctional	Specialized	Multifunctional	Specialized	Specialized	Multifunctional
Emulation accuracy	Acceptable	All right	All right	Acceptable	Excellent	All right	All right
Quality of collected data	Weak	Excellent	All right	All right	All right	All right	All right
Readability	Excellent	Excellent	All right	Excellent	All right	Acceptable	Acceptable
Scalability and performance	Excellent	All right	Acceptable	All right	Acceptable	Acceptable	Acceptable
Expandability	Excellent	Excellent	All right	Excellent	Excellent	Acceptable	All right
Embeddable	Acceptable	Excellent	Acceptable	All right	All right	Acceptable	All right
Complexity of setup and use	Acceptable	All right	All right	All right	All right	All right	All right
Maintenance requirements	All right	All right	Acceptable	All right	All right	All right	All right

Crowie may require significant time for individuals unfamiliar with honeypots to become proficient. Still, it is well-suited for researchers and analysts interested in studying attack patterns. According to this standard, Amun is considered satisfactory, and the only requirement to utilize the honeypot is to execute the primary Python script. Glasstopf is an excellent option, and the software can be obtained directly from the SVN repository. Kippo excels in this aspect, as its configuration is stored in a singular text file. On the other hand, it should be noted that installing Thug can be challenging and necessitates technical expertise.

Honeyd demonstrates exceptional scalability and performance. Specifically, its default configuration allows it to effortlessly manage 100 public, unutilized, and vulnerable IP addresses. Based on this criterion, Dionaea can monitor multiple network interfaces and IP addresses simultaneously. Cowrie has a flexible configuration that permits modifications without requiring a restart. According to this standard, Amun is considered high quality and capable of managing over 100 IP addresses on a single installation. Glasstopf is not inherently flawed, but it is important to note that the tool has limited performance capabilities. Based on this standard, Kippo possesses satisfactory attributes and can listen to numerous IP addresses. Thug can run multiple concurrent sessions on a single server, but its bandwidth is average.

Honeyd can be described as having a modular architecture in terms of its expandability. Dionaea excels in this aspect due to its modular architecture, which guarantees the ability to create customized extensions. Cowrie can be described as expandable and modular based on this criterion. Amun possesses exceptional extensibility and modular architecture like Glasstopf. According to this standard, Kippo is reliable, and Thug is satisfactory.

Regarding installation feasibility, it can be stated that Honeyd possesses satisfactory attributes and lacks any interface for intercommunication with other tools. The feasibility of installing Dionaea is exceptional, and all data is stored in the SQLite database. Cowrie is highly compatible with integration and is designed to operate seamlessly with well-known threat intelligence platforms and SIEM tools. According to this standard, Amun and Glasstopf are considered good, but Kippo lacks an API interface for seamless integration with other systems. Regarding Thug and this criterion, it is essential to highlight that the overall embeddability is satisfactory.

According to the maintenance requirement criterion for Honeyd, only minimal maintenance is needed to keep it valid. Based on this criterion, Dionaea exhibits minor issues, but there are no indications of potential future problems. Managing Cowrie is relatively intricate. Amun, Glasstopf, and Thug are all low-interaction honeypots that are relatively simple to maintain.

Specific honeypot solutions outperform others. For instance, Honeyd surpasses Dionaea in terms of buildability. Various downloadable versions of Honeyd can be found on the project's website. Dionaea relies on several third-party software prerequisites, each of which must be built

 doi:10.20944/preprints202408.0946.v1

separately. Automated generation was facilitated, and all requisite components were accessible at the time of authoring. Dionaea surpasses Honeyd in terms of emulated services. Dionaea is highly proficient in emulating services. The software replicates several vulnerable services, such as SMB, HTTP, FTP, TFTP, MSSQL, MySQL, and SIP, with high realism in their emulation. Honeyd is ineffective in simulating or imitating services. The simulation of listening services on the ports of the Honeyd monitor relies on Python scripts, with the most fundamental being included with a honeypot. The level of interaction these scripts offer is rudimentary, and achieving the most advanced emulation requires user implementation.

There is no superior honeypot to others; each has merits and drawbacks. The choice of honeypot depends on the specific goals you are aiming to accomplish. Honeypots are classified according to their level of interaction to understand their capabilities better. Increased engagement with the assailant leads to more excellent knowledge acquisition but also entails a heightened level of danger. BOF and Spectre are examples of low-interaction honeypots. They are simple to configure and entail minimal risk. Nevertheless, their capabilities are restricted to imitating specific services and operating systems primarily employed for detection. Mantrap and Honeynets are examples of honeypots that involve moderate to significant interaction.

Developing a honeypot system in cybersecurity entails constructing a regulated setting that appears appealing to potential intruders. A honeypot diverts attackers' attention from genuine systems while collecting data on their strategies, methods, and objectives. When designing a honeypot solution, it’s essential to take these processes into account:

- Selecting the type of honeypot - various types exist, each designed for distinct purposes. It is imperative to consider the kind most suits the organization's objectives. Several prevalent categories include:
    1. Low-interaction honeypots are emulators that mimic specific vulnerabilities or services, but their level of interaction is limited. They are simple to implement and maintain but provide limited information regarding the attackers;
    2. High-interaction honeypots imitate genuine systems and enable attackers to interact extensively. They collect comprehensive data but necessitate additional resources and maintenance;
    3. Virtualized Honeypots are implemented on virtual machines, which enables effortless deployment, isolation, and scalability;
    4. Manufacturing honeypots refers to natural systems or services that have been strengthened and closely observed. They can detect and prevent attacks on actual infrastructure while reducing the potential harm to operational systems.
- Network placement – it is imperative to determine the specific location within the network infrastructure where the honeypot system will be deployed. The primary emphasis should be on positioning the honeypot near actual systems to replicate their presence or isolating it within a distinct network segment to minimize the likelihood of the honeypot being utilized as a starting point for launching attacks on other systems.
- The choice of operating systems and services—The selection of the operating system and services that the honeypot system will imitate should be determined. To create a convincing honeypot, it is essential to replicate the systems being emulated closely. Verifying that the chosen systems and services possess well-documented vulnerabilities is imperative to attract attackers.
- Regularly updating and patching the honeypot system is necessary for effective vulnerability management and to ensure the presence of genuine vulnerabilities. This provides the honeypot's continued appeal to potential attackers and prevents it from becoming obsolete or easily recognizable.
- Robust monitoring and logging mechanisms are essential for recording all activities within the honeypot system, including network traffic, system logs, and user interactions. Observe deviant behavior and set up notifications for potential attacks or dubious actions.
- Deception techniques—Integrating deceptive strategies enhances the attractiveness of honeypots. This may involve presenting fabricated data or files, deploying honeytokens (counterfeit credentials or documents), or offering lure services designed to appear enticing to

 doi:10.20944/preprints202408.0946.v1

potential attackers. These factors enhance the likelihood of attackers contacting the honeypot and uncovering their techniques.

Additional security measures must be implemented to safeguard the honeypot system itself. These measures encompass separating it from vital systems, employing robust authentication methods, and implementing intrusion detection and prevention mechanisms. A customized incident response plan should be created for the honeypot system. This plan should provide a detailed plan of action to follow during an attack or compromise, including instructions on analyzing and responding to the collected data.

Thorough documentation and analysis are necessary for honeypot systems' design, implementation, and maintenance processes. Analyzing the gathered data regularly improves understanding of attacker strategies, weaknesses, and emerging dangers. This information is significant in enhancing the overall security of the network.

6. Best Practices and Recommendations

Our research's best practices and recommendations can be divided into three topics: honeypot implementation strategies, safety assumptions and mitigation techniques, and integration with existing security infrastructure. Let’s evaluate them now.

6.1. Honeypot Implementation Strategies

To select a strategy, companies should consider their susceptibilities, the specific types of attacks to which they are most susceptible, and their capacity to manage system maintenance. A low-interaction honeypot may be appropriate if a company is vulnerable to basic bot attacks.

There are multiple methods for implementing a honeypot system. The selection and quantity of honeypots will vary based on the desired objectives. Deploying a honeypot in any network location is ineffective and futile without sufficient understanding of the specific honeypot system and the necessary information. Before installation, it is essential to possess a fundamental understanding of the particular type of honeypot. Both technical and human resources significantly influence the implementation of a honeypot system. Therefore, successfully implementing high-interaction honeypot systems requires a deep understanding of computer science and expertise in security. While the primary objective of a honeypot system is to deceive the attacker and collect information about their activities, it is crucial to ensure that the attacker remains unaware of the presence of the honeypot system.

Implementing the honeypot system in a traditional configuration designed for online use involves making it accessible to the public network. This allows the system to observe attackers' actions and gather data on worms and other malicious behavior. The system's primary location is within the DMZ. A DMZ is a designated physical or virtual subnet that isolates an organization's specific services to a less trusted external network, such as the Internet.

Conversely, internal implementation gathers information about threats and infections within a specific organization's network. For instance, if one of the employees chooses to undermine a particular computer within the local network, the honeypot system deceives them during the engagement with the attacker, leading to the discovery of the attacker's behavior, identity, and tactics.

6.2. Security Assumptions and Mitigation Techniques

When an organization chooses to implement and utilize honeypots as a component of a multi-layered security defense, there are some factors to consider:

- The position of the honeypot is vital for its effectiveness. Strategically positioning honeypots within the network provides potential attackers with a target, as these honeypots appear to be legitimate services. The configuration should also consider the network environment where the attacker can establish a presence. Moreover, deploying honeypots on external or internal infrastructure will dictate the identified attacks' nature, frequency, and intensity. Outward-facing honeypots are vulnerable to regular scanning, deceptive botnet exploits, and practical

 doi:10.20944/preprints202408.0946.v1

attacks from various threat actors. This can amplify noise in logging signals, impeding distinguishing between genuine and targeted threats.

- Isolation—The honeypot must be segregated from the rest of the network and should not store confidential information. This reduces the likelihood of attackers using their access to honeypots to carry out further attacks across the network, using the honeypot as a starting point. Honeypot configurations should be designed to trap attackers and monitor their activities accurately and effectively.
- Tracking—The honeypot should be meticulously observed to collect data on the attacker's tactics, techniques, and procedures (TTPs) that can be incorporated into the network and security defenses. Mere implementation of a honeypot is insufficient. An essential task is to analyze honeypot logs to identify malicious activity, which serves to promptly notify the organization of possible attacks.
- Management and maintenance—The honeypot should be overseen by proficient security personnel who can set it up, monitor it, and uphold it to guarantee its efficacy in detecting and responding to attacks. Furthermore, it is crucial to regularly update the honeypot to maintain its authenticity as a replica of genuine systems and applications, thus ensuring its appeal to potential attackers.
- Integration with complementary security measures – The honeypot should be seamlessly incorporated into existing security measures, including firewalls, intrusion detection systems, incident response plans, SIEMs, MDRs, and systems for managing detection alerts, investigation, and response. This guarantees that the honeypot enhances the network's security and boosts its worth by generating activity detections and alerts. Honeypots can be utilized to dynamically construct detections by leveraging information from attackers.

Legal considerations should be considered when using honeypots, as certain countries have stringent regulations concerning the surveillance and interception of communications. Organizations must adhere to applicable laws and regulations when utilizing a honeypot and evaluate the potential risk of monitoring the wrong subjects.

Speed limit rules serve as a protective measure against denial-of-service (DoS) attacks. The system allows the administrator to permit or limit data movement from the origin to the destination. It guarantees that the server is not overwhelmed by excessive simultaneous requests. The system allows for the addition of items either individually, in groups, or for all members of the cluster simultaneously. The Security Manager is responsible for storing updates on intrusion signatures and conducting regular checks on the detection mechanism using exploit-based or anomaly-based intrusion methods to identify online or offline intrusions. This is the reason why the system promptly identifies any atypical traffic. Reference data, also known as data reference, is a term used to describe data used as a point of reference or comparison. It provides a standard or baseline against which other data can be measured or evaluated. The architecture incorporates reference data to store comprehensive information regarding recognized intrusion signatures or normal behavior profiles. This will establish the appropriate rhythm for additional intrusion and information acquisition, as the processing elements consistently refresh with new insights into the observed behavior. Configuration data, or config data, refers to the information that determines the settings and parameters of a system or application. The processing elements typically use the configuration data to store the intermediate results. Content filtering in the Content Filtering System (CFS) generally involves using software to restrict access to potentially harmful content.

6.3. Integration with Existing Security Infrastructure

There are three different approaches to integration. Implementation of honeypot technology within the network. This method entails placing the Honeypot in the network's LAN, DMZ, or Internet segment. By placing Honeypot systems within the LAN area of the network, these systems can be integrated into the security network. In this scenario, the Honeypot systems are in the same network segment as the production servers. The primary benefit of this design is that the Honeypot can identify a malevolent assault originating from both the Internet and the local network. A honeypot is a cybersecurity technique that attracts and deceives potential attackers to gather

 doi:10.20944/preprints202408.0946.v1

information about their methods and intentions. Figure 2 displays the placement of the Honeypot within the LAN segment of the network:

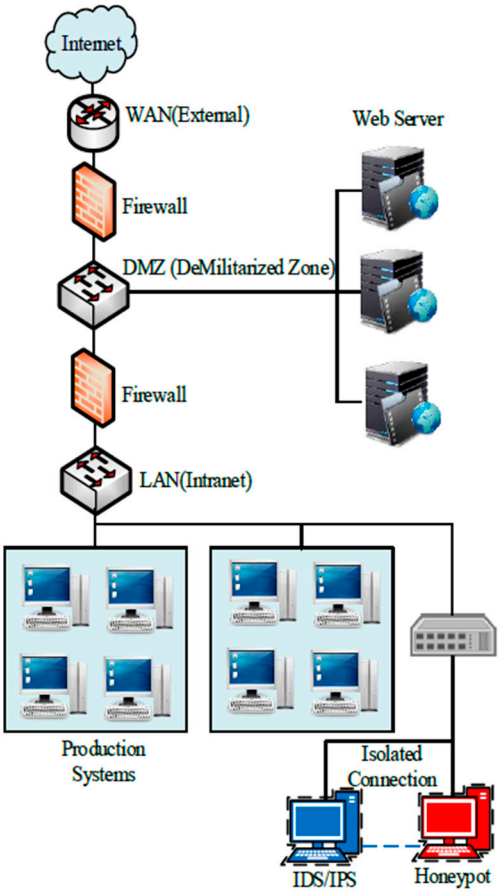

Figure 2. FLAX honeypot [23].

As depicted in Figure 2, a honeypot can identify external and internal malicious activities. It is widely recognized that a Honeypot in the LAN region poses substantial security risks. In this scenario, the attraction component of the attack Honeypot should have a low level of interaction. The primary benefit of placing Honeypots in the DMZ area is the segregation of the DMZ region from the local network. This scenario is not advisable as the sole security solution. The Honeypot in the DMZ cannot offer security for the Host in the local network, which is the cause of this issue. This particular Honeypot cannot detect any malicious attacks on the local network. It is suggested that an additional Honeypot within the local network be employed for this situation. This particular location is well-suited for detecting unauthorized individuals in the DMZ region. A DMZ region is a designated area that sometimes houses critical servers, such as web or mail servers [23]. Within this configuration. Activities related to Internet honeypots on the Honeypot platform. It is under surveillance by an Intrusion Detection System/Intrusion Prevention System. It is crucial to properly set up the Honeypot and establish network isolation between the IDS and the Honeypot. A honeypot is a cybersecurity technique that attracts and deceives potential attackers to gather information about their tactics and intentions. It lacks firewall protection due to its direct placement in the Internet region. Furthermore, the activation of Honeypot is prohibited to ensure the security of the LAN and DMZ areas. This approach primarily focuses on detecting external network attacks. This solution is deployed as a Honeypot on the Internet, outside the corporate network, as depicted in Figure 3:

 doi:10.20944/preprints202408.0946.v1

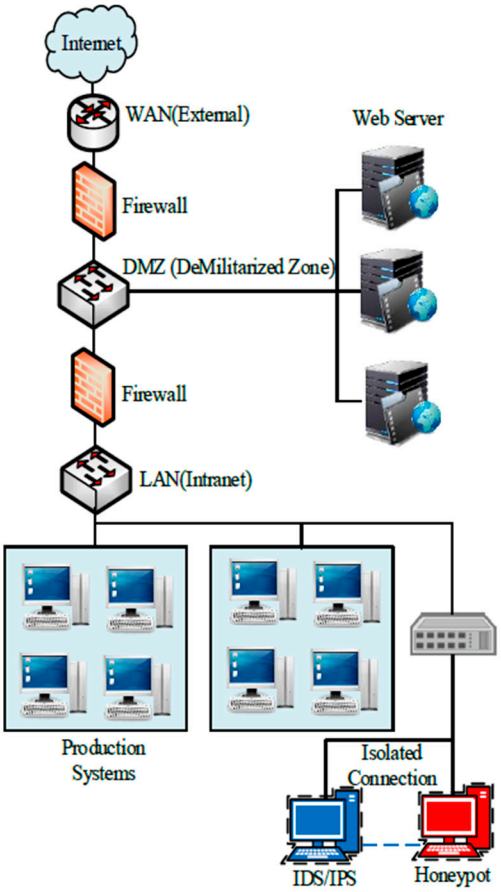

Figure 3. Internet honeypot [23].

Selecting these different approaches has pros and cons, and it's imperative to stay true to the best IT design principle—make the design fit the requirements. This section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation, and the experimental conclusions that can be drawn.

7. Future Work

With the increasing sophistication and evolution of cybersecurity threats, innovation has also made significant progress in developing honeypots. Multiple developmental trends in the field of honeypot innovation are expected to influence its future trajectory significantly. These trends include [24]:

- Progress in machine learning: Machine learning and artificial intelligence have been utilized to enhance the identification of attacks in honeypots. In the future, these advancements are expected to further progress, enabling more precise and automated detection of attacks.
- Enhanced automation is expected to impact the advancement of honeypots significantly. This will enable organizations to deploy honeypots more quickly and effectively, thereby improving their ability to analyze the data gathered by honeypots systematically.
- Honeypots are expected to be more effectively integrated with other security tools, such as intrusion detection and prevention systems, to enhance coordination and collaboration. By harnessing the attributes of different security devices, organizations can improve their ability to safeguard their systems against attacks.
- Cloud-based honeypots are expected to follow the trend of organizations migrating their infrastructure to the cloud. Cloud-based honeypots can distinguish between attacks on assets and attacks on cloud-based applications.
- Fraud technology, including honeypots, is a component of the broader field of cybersecurity. Fraud technology encompasses the dissemination of diversions or other deceptive methods employed to confuse attackers and create additional complications within the organization.

 doi:10.20944/preprints202408.0946.v1

These emerging patterns are expected to influence the advancement of honeypot technology significantly. Organizations can enhance their system security against cyberattacks by leveraging advancements in machine learning and artificial intelligence, increased mechanization, and integration with other security devices. Furthermore, advancements in cloud-based ports and misrouting will become progressively crucial as organizations migrate more of their infrastructure to the cloud and attackers persist in enhancing their tactics. The advancement of honeypot technology is expected to remain an essential tool in combating cybercrime. These trends in development will ensure its continued success in addressing current challenges.

8. Conclusions

Cyber threats are becoming more advanced and specifically designed to target individuals or organizations to exploit weaknesses. Hence, individuals and organizations must combat malware and cyberattacks proactively. This paper discusses the integration of honeypots into the overall security system to enhance existing security measures, including firewalls, intrusion detection systems, and incident response protocols. Implementing best practices and procedures, such as turning off unnecessary or insecure services, addressing necessary remediations, and employing robust authentication mechanisms, is essential for deterring malicious individuals.

Honeypots are practical tools for gathering accurate data on the progression of threats. Various honeypot solutions are employed to address specific types of threats. Research honeypots can be utilized to collect data and analyze the methods used in malicious activities. Production honeypots seamlessly incorporate into pre-existing security systems to enhance the capacity to identify and respond to attacks. Honeypots offer several benefits, including gathering limited data that may contain valuable information. Honeypots can be detected using minimal resources, and even low-spec systems are sufficient to operate them. They can operate in encrypted or IPv6 systems or environments. They are highly uncomplicated and adaptable to use.

The data obtained from the honeypot can offer valuable insights into attack patterns, classifications, and the modus operandi of assailants. The data can be utilized to enforce new Intrusion Detection System (IDS) regulations, pinpoint vulnerabilities of attackers with an interest in the system, ascertain the organization of attacks, and determine the attack models employed. Upon detecting a successful attack, the organization must possess the capability to respond promptly. The most frequently discussed legal concerns regarding honeypots are privacy (the potential violation of hackers' privacy), liability, and capture/trapping.

After evaluating the selected honeypot solutions, namely Honeyd, Dionaea, Cowrie, Amun, Glasstopf, Kippo, and Thug, it is evident that some honeypots outperform others. The choice and implementation of each honeypot solution should be tailored to the organization's specific requirements. Practically, various development patterns in the innovation of honeypots are expected to impact this field significantly. These include advancements in machine learning and artificial intelligence, enhanced automation, improved integration of honeypots with other security tools, cloud-based honeypots, and the application of fraud technology.

The research involved testing different honeypot tools to identify potential vulnerabilities and analyze attacks. The network attacks were simulated using tools such as nmap and Metasploit. The findings unequivocally demonstrated that Amun, Dionaea, Cowrie, and Thug efficiently identified and logged various attacks, encompassing network scans, connection attempts, and command execution. However, Glasstopf and Honeyd could not be initiated due to problems with installation and configuration, resulting in a lack of activity tracking. Each honeypot tool possesses unique advantages and limitations, allowing it to cater to the specific requirements and objectives of the user. Cowrie is notable for its straightforward installation process, seamless integration with Python, and capability to replicate various protocols. Dionaea is noteworthy for its extensive configuration options, compatibility with multiple architectures, and ability to detect and intercept a wide range of attacks. Although it can detect Amun, it has restricted functionality. The thug has meticulously recorded malevolent activities originating from websites, furnishing invaluable data that is pivotal in identifying and mitigating potential hazards to the local system. Upon examining the analysis, it

 doi:10.20944/preprints202408.0946.v1

can be inferred that each honeypot possesses unique characteristics, and the selection of an appropriate solution for the organization depends on the complexity it is willing to embrace and the organization's requirements.

When selecting a strategy, companies should consider their susceptibilities, the specific types of attacks they are most prone to, and their capacity for system maintenance. The two most prevalent deployment strategies are the internet-oriented classic installation and the internal implementation to gather information about threats and infections within a specific organization's network. Several safety assumptions and honeypot mitigation techniques are commonly used. Several mitigation techniques include honeypot detection, host-based intrusion detection systems, network-based intrusion detection systems, geo-routing, employing a straightforward defense technique with MAC procedure, and implementing a puzzle technique.

Author Contributions: Conceptualization, Z.M., and L.M.; methodology, Z.M., and Z.K.; software, Z.M.; validation, Z.K., and G.D.; formal analysis, Z.M., and G.D.; investigation, Z.K., and G.D.; resources, L.M.; data curation, Z.M., and L.M.; writing—original draft preparation, Z.K., and G.D.; writing—review and editing, L.M., and Z.M.; supervision, L.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data supporting this study's findings are available from the corresponding author, [L.M.], upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Šimon M, Huraj L, Hrinkino D. Using a honeypot to improve student cybersecurity awareness. 2023 21st International Conference on Emerging eLearning Technologies and Applications (ICETA). 2023 Oct 26; doi:10.1109/iceta61311.2023.10343633
2. Nintsiou M, Grigoriou E, Karypidis PA, Saoulidis T, Fountoukidis E, Sarigiannidis P. Threat intelligence using digital twin honeypots in cybersecurity. 2023 IEEE International Conference on Cyber Security and Resilience (CSR). 2023 Jul 31; doi:10.1109/csr57506.2023.10224997
3. Touch S, Colin J-N. Asguard: Adaptive self-guarded honeypot. Proceedings of the 17th International Conference on Web Information Systems and Technologies. 2021; doi:10.5220/0010719100003058
4. Subhan D, Lim C. Analyzing Adversary's Attack on Ethereum Collected from Honeypots. 2023 11th International Conference on Information and Communication Technology (ICoICT). 2023:313-318. DOI: 10.1109/ICoICT58202.2023.10262563.
5. Benedict S. EA-pot: An explainable AI-assisted Blockchain Framework for honeypot IP predictions. Acta Cybernetica. 2022 Nov 22;26(2):149–73. doi:10.14232/actacyb.293319
6. Valeros V, Rigaki M, García S. Attacker Profiling Through Analysis of Attack Patterns in Geographically Distributed Honeypots. arXiv. 2023. DOI: 10.48550/arXiv.2305.01346.
7. Baykara M, Das R. SoftSwitch: a centralized honeypot-based security approach using software-defined switching for secure management of VLAN networks. Turk J Electr Eng Comput Sci. 2019;27:3309-3325. DOI: 10.3906/elk-1812-86.
8. Du C, Zhao S, Wang W. RRPOT: A Record and Replay Based Honeypot System. J Phys Conf Ser. 2021;1757:012183. DOI: 10.1088/1742-6596/1757/1/012183.Naik N, Shang C, Jenkins P, Shen Q. D-FRI-Honeypot: A Secure Sting Operation for Hacking the Hackers Using Dynamic Fuzzy Rule Interpolation. IEEE Trans Emerg Top Comput Intell. 2020;5:893-907. DOI: 10.1109/TETCI.2020.3023447.
9. Schuba M, Höfken H, Linzbach S. An ICS Honeynet for Detecting and Analyzing Cyberattacks in Industrial Plants. 2021 International Conference on Electrical, Computer and Energy Technologies (ICECET). 2021:1-6. DOI: 10.1109/ICECET52533.2021.9698746.
10. Nilă C, Preda M, Apostol I, Patriciu V. Reactive WiFi honeypot. 2021 13th International Conference on Electronics, Computers and Artificial Intelligence (ECAI). 2021:1-6. DOI: 10.1109/ECAI52376.2021.9515048.
11. Draghicescu D, Caranica A, Fratu O. Honeypot Technologies for Malware Detection and Analysis. Strategies XXI - Command and Staff College. 2021. DOI: 10.53477/2668-2028-21-34.
12. Sehgal R, Majithia N, Singh S, Sharma S, Mukhopadhyay S, Handa A. Honeypot Deployment Experience at IIT Kanpur. 2020:49-63. DOI: 10.1007/978-981-15-1675-7_6.

 doi:10.20944/preprints202408.0946.v1

13. Sladić M, Valeros V, Catania C, Garcia S. LLM in the Shell: Generative Honeypots. arXiv. 2023. DOI: 10.48550/arXiv.2309.00155.
14. Subhan D, Lim C. Unveiling Attack Patterns: A Study of Adversary Behavior from Honeypot Data. 2023 IEEE International Conference on Cryptography, Informatics, and Cybersecurity (ICoCICs). 2023:178-183. DOI: 10.1109/ICoCICs58778.2023.10276516.
15. Florea R, Craus M. A Game-Theoretic Approach for Network Security Using Honeypots. Future Internet. 2022;14:362. DOI: 10.3390/fi14120362.
16. Mocanu F, Scripcariu L. Intrusion Detection Platform with Virtual Honeypots. 2023 International Symposium on Signals, Circuits and Systems (ISSCS). 2023:1-4. DOI: 10.1109/ISSCS58449.2023.10190854.
17. Katakwar H, Aggarwal P, Dutt V. Modeling the effects of different honeypot proportions in a deception-based security game. Human Factors Cybersecurity. 2023. DOI: 10.54941/ahfe1003727.
18. Ziaie Tabari A, Ou X. A Multi-phased Multi-faceted IoT Honeypot Ecosystem. Proc 2020 ACM SIGSAC Conf Comput Commun Secur. 2020. DOI: 10.1145/3372297.3420023.
19. Osman M, Nadeem T, Hemida A, Kamhoua CA. Optimizing Honeypot Placement Strategies with Graph Neural Networks for Enhanced Resilience via Cyber Deception. Proc 2nd Graph Neural Netw Workshop. 2023. DOI: 10.1145/3630049.3630169.
20. European Network and Information Security Agency (ENISA). Proactive Detection of Security Incidents. 2012. Available from: https://www.enisa.europa.eu/publications/proactive-detection-of-security-incidents-II-honeypots/@@download/fullReport, accessed May 2024.
21. Yang X, Yuan J, Yang H, Kong Y, Zhang H, Zhao J. A Highly Interactive Honeypot-Based Approach to Network Threat Management. 2023. Available from: https://www.mdpi.com/1999-5903/15/4/127, accessed November 2023.
22. Baykara M, Das R. A Survey on Potential Applications of Honeypot Technology in Intrusion Detection Systems. 2015. Available from: https://www.researchgate.net/publication/290433369_A_Survey_on_Potential_Applications_of_Honeypot_Technology_in_Intrusion_Detection_Systems, accessed May 2024.
23. Satheesh V. The Evolution Of Honeypot Technologies And Future Directions. Available from: https://www.irjmets.com/uploadedfiles/paper/issue_5_may_2023/38648/final/fin_irjmets1683961611.pdf, accessed May 2024.