10.53477/2668-2028-21-34

HONEYPOT TECHNOLOGIES FOR MALWARE DETECTION AND ANALYSIS

Dragoș DRĂGHICESCU, PhD
Research staff, Faculty of Electronics, Telecommunications and Information Technology,
University Politehnica of Bucharest
dragos.draghicescu@etti.pub.ro

Alexandru CARANICA, PhD
Research staff, Faculty of Electronics, Telecommunications and Information Technology,
University Politehnica of Bucharest
alexandru.caranica@speed.pub.ro

Octavian FRATU, PhD
Professor, Faculty of Electronics, Telecommunications and Information Technology,
University Politehnica of Bucharest
ofratu@elcom.pub.ro

Abstract*: In this paper, we offer a brief summary of latest developments in honeypot technologies, used for malware detection and analysis. This includes not only honeypot software, but also methodologies to analyze captured honeypot data. As such, our focus in this work is to keep track of current developments related to traffic analysis, especially honeypot technologies, as a means of data capture and interpretation of malicious traffic. Zero-day attacks are still very hard to predict, then handle, by any security platform. Means to successfully predict an attack is of paramount importance to the world of cybersecurity. Effective network security administration depends, to a great extent, on the understanding of existing and emerging threats propagated over the web. In order to protect information systems and its users, it is of crucial importance to collect accurate, concise, high-quality information about malicious activities, for security researchers to be able to reverse-engineer, then understand and stop a malicious actor.*

Keywords: *computer security; honeypot technologies; malware analysis.*

Introduction

The primary goal of antimalware systems is to defend computers against attacks launched by malicious users. In recent years, intrusions and attacks through the Internet have increased notably[1]. This increase in the number of incidents has been accompanied by a clear evolution of the tools and techniques used by the aggressors. People that are responsible about these aggressions are usually very talented in regards to software development and able to develop tools and software that permit them to revisit a compromised target, being using a backdoor or a direct connection. The attacker behavior can be very complex, and specialists thought about several ways to solve the problem of continuous cyber-threats. The need to understand the attacker and its methods turned into a class of tools named “honeypots”. The concept of a “honeypot system” older than two decades[2]. A honeypot is a deceptive mechanism that presents itself as an open device (or software), in order to lure in potential attackers and / or exploit actors. The presence of a potential attacker inside a honeypotting system permits the security specialist to observe all its actions by logging, collecting information and analysis. Its purpose can be to either figure out the pattern of the attack or find a way to stop the attack’s spreading or close the breech exploited by the intruded malicious software.

These systems are paramount in designing the latest tools and methods for computer / network security detection and prevention. Honeypots can capture and log a great number of behavioral patterns or uncover spreading techniques used by attackers “in-the-wild”. Results gathered can be used to either improve the current security of the network, or for research purposes, such as learning about new exploits, depending on the nature of the honeypot.

[1] DV Silva, GDR Rafael, “A review of the current state of Honeynet architectures and tools”, *International Journal of Security and Networks 12 (4)*, 255-272, 2017.

[2] Anisi Behnam, *State-of-the-art Evaluation of Low and Medium Interaction honeypots for Malware Collection*, Degree Thesis, Edinburgh Napier University, School of Computing, Aug 2016.

As malware methods progressed, so did the honeypot-based tools to protect against attacks, that most organizations or personal users face[3]. We iterate, below, some of these commonly used tools, for network-connected devices, that nowadays are the major targets for malicious actors.

One of the most used tools for network protection are firewalls, which help protect inside devices from outside attacks. IDS (Intrusion Detection Systems)[4] are used to complement firewalls, enabling network administrators to protect devices from direct exposure to the outside WAN (Wide Area Network). These "essential" tools provide a useful "first encounter" tool to protect devices from potential tons of initial attacks, like scanner tools, device discovery, etc. But these tools sometimes lack the ability to detect new threats like zero-day vulnerabilities, based on newly discovered exploits[5]. They also fail to collect more information about the attacker's malicious future intent, like a potential payload deployment. For example, signature based IDSs or antiviruses are not capable of detecting these zero-day unknown attacks, because they do not have the signatures for these new attacks in their database. This is where honeypots come into play, as early catching an attackers exploit/payload by a security research organization, allows signature-based, or behavioral systems to analyze the malwares sample at execution time and act, before the malware spreads through the network.

This generalized idea, often mythicized on the profile of the intruders, makes people believe that only those devices containing important or classified information will be subject to a cybernetic attack. The selective attacks directed by experts are a small percentage of the attacks that take place every day through the web[6]. Most incidents that occur on the internet are not aimed at computing devices or specific organizations/individuals. Instead, their objective is the easy victim, in order to extract immediate benefits, like including the infected device in a spam network or, like recent events have shown, blackmail recipients for decrypting their device, after a crypto-malware attack. The target selected can be any device that is connected and has a specific weakness (software bug) that the aggressor is looking for and can take advantage of to gain access to the system (by exploiting the found vulnerability). At present, it is not necessary to have in-depth knowledge of the functioning of a system to be able to attack it. In fact, most intruders are limited to using tools created by others, many of which can be directly downloaded from internet and are simple to use[7].

Background and related work

In[8], the author makes a brief history of "honeypots" systems: during the initial decade of the 21st century, there was a gradual increase in both Linux and Windows-based malware/worms and as recently, even "Apple Silicon" is being targeted more often, as popularity of these devices is on the rise[9]. These malicious files, historically, have proved to be extremely effective in taking down production systems. "One of the challenges that various security researchers faced was obtaining a copy of these malicious file for further detailed analysis. It was difficult to get copies from already infected systems, because of data tampering, or because these malicious files resided only in the system's memory. At that time, honeypots proved to be a powerful tool that can capture these malware actors without harming any production system. Fred Cohen's Deception ToolKit in 1998 was one of the first publicly

[3] Jeffrey Nkwetta Asonganyi, "Honey-System: Design, Implementation and Attack Analysis", College of Technology, University of Buea, in partial fulfilment of the requirements for the award of the Degree of Bachelor of Technology, 2018.

[4] *Intrusion Detection System*, Wikipedia.org, https://en.wikipedia.org/wiki/Intrusion_ detection_ system, accessed of February 2021.

[5] "Zero-day (computing)," Wikipedia.org, https://en.wikipedia.org/wiki/Zeroday_ (computing), accessed February 2021.

[6] Silva Dv, Rafael Gdr, *op. cit.*

[7] *Idem.*

[8] Jeffrey Nkwetta AsonganyI, *op. cit.*

[9] Carly Page, "The first M1 MacBook malware has arrived – here's what you need to know", TechRadar https://www.techradar.com/news/the-first-m1-macbook-malware-has-arrived-heres-what-you-need-to-know, accessed of February 2021.

available honeypot. It was intended to make it appear to attackers as if the system running DTK had a large number of widely known vulnerabilities", according to[10].

The security of the data in early times was exclusively defensive, trying to keep the intruders away. Since then, techniques, tools and technologies have been developed to stop attacks. In 1992, Marcus Ranum developed firewalls in his release TIS Firewall Toolkit. This collection of code was used as one of the first controls for public nets of access by catwalks with the aim of avoiding any type of unauthorized entry[11]. At that time, security professionals and researchers made attempts to learn about these types of attack. However, those attempts were limited in effort and reachability because a great deal of the information obtained and published was based on the technical details of the exploits used by invaders, stressing the vulnerabilities of the objective and the explanation of how the exploit was taking advantage of that vulnerability.

In October 1999, a group of individuals that sought to learn more about attacks, threats and vulnerabilities was formed by Marty Roesch (developer of the system of detection of intruders called Snort[12]), Cris Brenton, J.D Glazer, Ed Skoudis and Lance Spitzner (author of The Honeynet Project[13]), the self-titled Wargames mail list. They worked in the construction of computers that were used to damage others, developing attack capabilities and methodologies of analysis to understand how they had been attacked. This group grew and became what is now known as The Honeynet Project[14].

But modern-day companies require safety solutions that will not only impede the access of the most advanced intruders, but also allow someone to study their behavior[15]. A Honeypot`s value consists in being able to also monitor malware samples and attacks[16]. Honeypots can be attacked or engaged according to a level of interaction. They can be classified as low, medium or high interaction[17].

Normally, Honeypots of low interaction work exclusively by emulating operating systems and services. The aggressor's activities are limited to the level and quality of the emulation of the Honeypot[18].

Honeypots of medium interaction are slightly more sophisticated than Honeypots of low interaction. Honeypots of medium interaction provide the aggressor a better illusion of an operating system, so that the aggressor can interact with. Therefore, more complex attacks can be registered and analyzed[19].

High interaction honeypots constitute a complex solution because they imply the use of operating systems and real applications, implemented in real hardware, without the need of using software emulation. It is important also to mention some of the most popular types of attack vectors that malware authors use today[20]:

[10] *Honeypot (computing)*, Wikipedia.org, https://en.wikipedia.org/wiki/Honeypot_ (computing), accessed of February 2021

[11] Silva Dv, Rafael Gdr, *op. cit.*

[12] Z. Zhou, Z. Chen, T. Zhou and X. Guan, "The study on network intrusion detection system of Snort," 2010 International Conference on Networking and Digital Society, Wenzhou, China, 2010, doi: 10.1109/ICND S.2010.5479341.

[13] Honeynet Project, Wikipedia: https://en.wikipedia.org/wiki/Honeynet_Project, accessed February 2021.

[14] *Idem.*

[15] I. Kuwatly, M. Sraj, Z. Al Masri and H. Artail, *A dynamic honeypot design for intrusion detection*, In Pervasive Services, ICPS 2004, IEEE/ACS International Conference on IEEE, 2004.

[16] L. Spitzner, *The honeynet project: Trapping the hackers*, IEEE Security&Privacy, 1(2), 2003, pp. 15-23.

[17] I. Mokube and M. Adams, *Honeypots: concepts, approaches, and challenges*, In Proceedings of the 45th annual southeast regional conference, ACM, 2007, pp. 321-326.

[18] A. Mairh, D. Barik, K. Verma and D. Jena, *Honeypot in network security: a survey*, In Proceedings of the 2011 international conference on communication, computing&security, ACM, 2011.

[19] P. S. Huang, C. H. Yang and T. N. Ahn, *Design and implementation of a distributed early warning system combined with intrusion detection system and honeypot*, In Proceedings of the 2009 International Conference on Hybrid Information Technology, ACM, 2009.

[20] N. Ansona, Dr. S. Sasidhar Babu, M. Sheema, Prof. P. Jayakumar, *Integrated Honeypot*, IJCET, 2014.

- **“Keylogger-Spyware Attack**: Spyware is a broad category of software designed to intercept or take partial control of a computer's operation without the informed consent of that machine's owner or legitimate user. In simpler terms, spyware is a type of program that watches what users do with their computer and then sends that information over the internet. Spyware or Keylogger can collect many different types of information about a user records the types of websites a user visits, records what is typed by the user to intercept passwords or credit card numbers, used to launch “pop up” advertisements. A key-logger spyware contains both scripts key-logger and spyware in a single program. These types of attacks can be very dangerous, as they can record keystrokes and potential banking passwords, for example, then send them to a malicious actor”[21].
- **“SSH attacks**: Nowadays, malicious users can easily find internet-faced servers that can be easily exploited and used for their activities. One of the most vulnerable targets in servers is the Secure Shell (SSH), used for remote access. Multiple times these servers got exploited by the hackers if a very weak password is used as an authentication mechanism. Whenever the hacker finds a device with an SSH service, he will try to “brute-force” it`s way in, in order to get unauthorized access. If the hacker succeeds, he gains remote access to the machine and then he can use it for malicious activities”[22].
- **Zero-Day attacks**: A zero-day attack or threat is an attack that exploits a previously unknown vulnerability in a computer application, meaning that the attack occurs on “day zero” of awareness of the vulnerability. This means that the developers have had zero days to address and patch the vulnerability. Malware writers can exploit zero-day vulnerabilities through several different attack vectors. Web browsers are a particular target because of their widespread distribution and usage. Attackers can also send e-mail attachments, which exploit vulnerabilities in the application opening the attachment. A special type of vulnerability management process focuses on finding and eliminating zero-day weaknesses. This unknown vulnerability management lifecycle is a security and quality assurance process that aims to ensure the security and robustness of both in-house and third-party software products by finding and fixing unknown (zero-day) vulnerabilities. The unknown vulnerability management process consists of four phases: a) analyze b) test c) report and d) mitigate”[23].

Usage of honeypots

Honeypots can be classified by their purpose in research and production. Research honeypots have the purpose of informing a research institution or anti-virus laboratory about general malicious intentions and productions honeypots intend to observe, analyze, catch and repel any attack activity of a specific network regarding its own traffic.

One of the most well knows principles in cyber security implementations is AHAT, meaning “always have an audit trail”. It suits well when trying to know what went wrong in case of an attack and works hand in hand with law regulations.

Many of today’s security tools and technologies are based on detection and alerting. Therefore, it is very common that security professionals tend to be overwhelmed with red flags risen by security equipment. When there is a lot of important information to look at, it all becomes irrelevant and real attacks are very hard to spot. In this situation the security practices will greatly benefit from the concept of a honeypot.

Honeypot systems present advantages in monitoring high productivity systems, but the risk should be permanently evaluated. A poor setup can facilitate attackers in using a pivotal point in launching an external or internal network attack. Honeypots used for the evaluation of internal threats can also be subject of legitimate user access, thus making it harder to spot disgruntled employees trying to abuse in some way the company internal resources. Having

[21] *Idem.*
[22] *Idem.*
[23] *Idem.*

knowledge of the deceptive technology techniques[24] helps increase the ratio between illegitimate user traffic and the legitimate one.

Incident handling in enterprise environments is reactive and proactive in nature and mean preparation, identification, containment, eradication, recovery and learning.

In the case honeypots are used for malware forensics investigations the objectives of security team are:

- Gathering information about the aim of a program;
- Gathering information about the mechanisms of an attack;
- Gather information about how an attacker software influences and changes the host;
- Gather information about how an attacker software makes changes to a program;
- Gather information about an attack complexity and its procedures;
- Gather information about the attack spreading and the damage it caused.

Most of the case studies presented in white papers for malware analysis present some form of honeypot used for capturing malicious intended traffic. As the intended weak system is designed to engage only with malicious traffic, all captured packets can be tagged as malicious. This approach has great benefits for designing behavioral tracking systems that perform excellent against "zero-day" attacks[25]. In Figure 1, we present a general scheme for a honeypot usage.

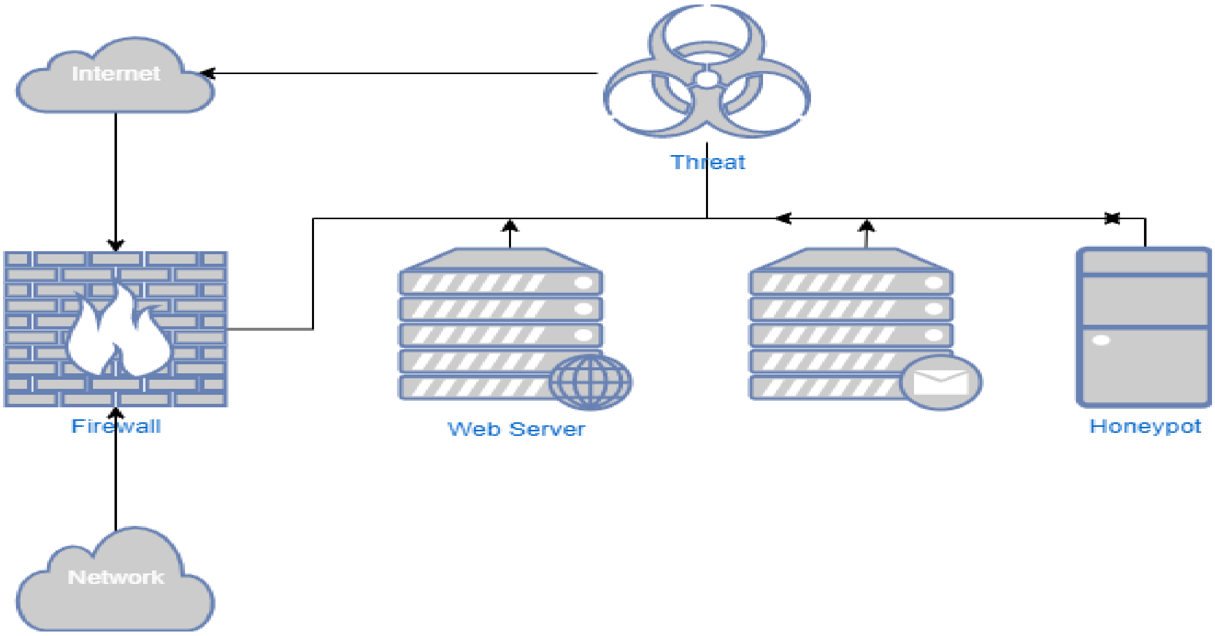

Figure 5. Honeypot usage example

Conclusions

Honeypots strong points can be summarized[26] in:

- Valuable data collection capability – the honeypot is an important source for traffic analysis;
- Independent from workload – the honeypot is not a production system;
- Zero-day-exploit detection – unknown malicious software can be captured and analyzed;
- High confidence traffic tagging – the traffic to and from the honeypot can be safely tagged as potentially malicious;

[24] Linan Huang, Quanyan Zhu, *Game of Duplicity: A Proactive Automated Defense Mechanism by Deception Design*. arXiv:2006.07942 [cs.GT], 14.03.2020.

[25] Oluwashola David Adeniji, Oluwadare Oluwasola Olatunji, *Zero Day Attack Prediction with Parameter Setting Using Bi Direction Recurrent Neural Network in Cyber Security*. International Journal of Computer Science and Information Security (IJCSIS), Vol. 18, No. 3, March 2020.

[26] Marcin Nawrocki, Matthias Wahlisch, Thomas C. Schmidt, Christian Keil, Jochen Schonfelder, *A Survey on Honeypot Software and Data Analysis*, 22 August 2016.

- Modularity – capability to be adjusted for different tasks.

In contrast, the problems that rise from using honeypots are:

- Narrow perspective – honeypots must receive traffic in order to be efficient;
- Increasing risks of self-infection if poorly managed;
- The possibility of identification or fingerprinting – if an attacker manages to identify that a honeypot is in use based on certain characteristics or specific behaviors, its value is nullified.

Acknowledgements

This research work was supported by a grant on the Romanian Ministry of Innovation and Research, UEFISCDI, project number 8SOL/2018 within PNCDIII, project code: PNIII-P2-2.1-SOL-2017-09-0102, project name: Integrated Information System for Management of Activities (IISMA) (http://siima.pub.ro/en/home/).

BIBLIOGRAPHY

1. SILVA D.V., RAFAEL G.D.R., “A review of the current state of Honeynet architectures and tools”, *International Journal of Security and Networks 12 (4)*, 255-272, 2017.
2. BEHNAM Anisi, “State-of-the-art Evaluation of Low and Medium Interaction honeypots for Malware Collection”, *Degree Thesis*, Edinburgh Napier University, School of Computing, Aug 2016.
3. ASONGANYI Jeffrey Nkwetta, “Honey-System: Design, Implementation and Attack Analysis”, College of Technology, University of Buea, in partial fulfilment of the requirements for the award of the Degree of Bachelor of Technology, 2018.
4. “Intrusion Detection System,” Wikipedia.org, https://en.wikipedia.org/wiki/Intrusion_ detection_ system, accessed February 2021.
5. “Zero-day (computing),” Wikipedia.org, https://en.wikipedia.org/wiki/Zeroday_ (computing), accessed February 2021.
6. PAGE Carly, “The first M1 MacBook malware has arrived – here's what you need to know”, TechRadar https://www.techradar.com/news/the-first-m1-macbook-malware-has-arrived-heres-what-you-need-to-know, accessed February 2021.
7. “Honeypot (computing),” Wikipedia.org, https://en.wikipedia.org/wiki/Honeypot_ (computing), accessed February 2021.
8. ZHOU Z., CHEN Z., ZHOU T. and GUAN X., "The study on network intrusion detection system of Snort," 2010 International Conference on Networking and Digital Society, Wenzhou, China, 2010, doi: 10.1109/ICNDS.2010.5479341.
9. *Honeynet Project*, Wikipedia: https://en.wikipedia.org/wiki/Honeynet_Project, accessed February 2021.
10. KUWATLY I., SRAJ M., AL MASRI Z. and ARTAIL H., *A dynamic honeypot design for intrusion detection*, In Pervasive Services, ICPS 2004, IEEE/ACS International Conference on IEEE, 2004
11. SPITZNER L., *The honeynet project: Trapping the hackers*, IEEE Security&Privacy, 1(2), 2003.
12. MOKUBE I. and ADAMS M., *Honeypots: concepts, approaches, and challenges*, In Proceedings of the 45th annual southeast regional conference. ACM, 2007.
13. MAIRH A., BARIK D., VERMA K. and JENA D., *Honeypot in network security: a survey*, In Proceedings of the 2011 international conference on communication, computing&security, ACM, 2011.
14. HUANG P. S., YANG C. H. and AHN T. N., *Design and implementation of a distributed early warning system combined with intrusion detection system and honeypot*, In Proceedings of the 2009 International Conference on Hybrid Information Technology, ACM, 2009.
15. ANSONA N, Dr. BABU S. Sasidhar, SHEEMA M., Prof. JAYAKUMAR P., *Integrated Honeypot*, IJCET, 2014.

16. HUANG Linan, ZHU Quanyan, *Game of Duplicity: A Proactive Automated Defense Mechanism by Deception Design*. arXiv:2006.07942 [cs.GT], 2020-06-14.
17. ADENIJI Oluwashola David, OLATUNJI Oluwadare Oluwasola, *Zero Day Attack Prediction with Parameter Setting Using Bi Direction Recurrent Neural Network in Cyber Security*. International Journal of Computer Science and Information Security (IJCSIS), Vol. 18, No. 3, March 2020.
18. NAWROCKI Marcin, WAHLISCH Matthias, SCHMIDT C. Thomas, KEIL Christian, SCHONFELDER Jochen, *A Survey on Honeypot Software and Data Analysis*, 22 August 2016.