Honeypots for Internet of Things Research: An Effective Mitigation Tool

Shen Xin En[1] Liu Si Ling[1] Fan Cheng Hao[1]

[1]Technische Universität München Asia
{shenen ,liusiing,fanchhao}@tum.edu.sg

Abstract

In recent years, due to their frequent use and widespread use, IoT (Internet of Things) devices have become an attractive target for hackers. As a result of their limited network resources and complex operating systems, they are vulnerable to attacks. Using a honeypot can, therefore, be a very effective way of detecting malicious requests and capturing samples of exploits. The purpose of this article is to introduce honeypots, the rise of IoT devices, and how they can be exploited by attackers. Various honeypot ecosystems will be investigated further for capturing and analyzing information from attacks against these IoT devices. As well as how to leverage proactive strategies in terms of IoT security, it will provide insights on the attack vectors present in most IoT systems, along with understanding attack patterns.

Keywords: Internet of Things, HoneyPot, Cyber-Attacks

1 Introduction

It's no secret that technology has become increasingly interconnected over the last few years. These types of devices are referred to as Internet of Things (IoT) and offer a variety of services to users. Wearable technology (i.e. smart watches) and smart devices (e.g. fridges, microwaves, cameras, etc.) fall into this category. In addition to reducing workloads and automating many tasks, they will increase quality of life for the users. [1–3] predicted by the year 2015 that there will be 20.4 billion connected devices installed in the world. While these devices have become increasingly connected, there is the potential for massive data breaches as a result of this increasing connectivity. There has been an increase in attacks against IoT devices since they have become more commonplace. There is a possibility that the Internet of Things could alter the way in which an hacker behaves and reacts in the cyber-attack landscape. In addition to providing communications, schedules, payments, and more, these devices also contain potentially more sensitive information about a person's personal life. The consequences of data breaches might be far more damaging as a result. Even cameras which are connected to WiFi can be hacked. This allows attackers to have the ultimate invasion of privacy as well as to obtain other essential details which can put a person's life in danger [4]. IoT devices are also being pushed to the market without much regard to security precautions in the tech industry, which favors the market release over security precautions. In the relatively new IoT ecosystem, many people just assume that companies wouldn't release any devices that potentially exposed their information to further attacks, however that isn't always true. There is a risk of exploits and hacks occurring if certain vulnerabilities are not considered. For example, if default credentials are used, or ports are left open, this could lead to easy hacks and exploits. The number of attacks has been increasing within the past couple years, as we mentioned before. There has been an increase in sophistication of the attacks as their vectors have been continuously changed. In spite of the fact that IoT attacks have increased in number, very few studies have been conducted to determine how effective these threats are and how broad their scope is. In order to patch the vulnerabilities in IoT devices, researchers need to understand the reasons and methods by which these attackers attack the devices. In this case, honeypots have proven to be an invaluable tool in the investigation process.

2 Honeypot Background

2.1 Honeypot Definition

In order to be successful, honeypots are designed with the sole intent of enticing attackers to compromise a dummy system. Devices such as these are configured to work within an isolated and separate network so they appear to be networked devices. Additionally, honeypots do not in fact provide any useful services to organizations directly, so having access to them can be seen as an act of malicious intent. The honeypot collects specific information from the attempts of an attacker residing on the honeypot in order to identify the attacker. The purpose of this technology is to identify weaknesses in a system and to uncover the tools and techniques that were used to compromise it in order to mitigate and prevent future attacks.

2.2 Honeypot Category and Classification

It is possible to classify and categorize honeypots in several different ways. As a starting point, they can be categorised according to how much they interact and what their purpose is. According to their purpose, honeypots have two uses: either for research or for production. “A production honeypot is used for protecting a company, whereas a research honeypot is used to learn more about the company” [5]. Most often, production honeypots are used in an organization’s environment to help minimize the risks that an attacker may pose to that organization. Thus, they are capable of detecting attacks, and they are much easier to build and configure because they have fewer features to contend with. Honeypots, which are used as a part of production environments, have the benefit of preventing false positives such as those that exist in traditional intrusion detection systems [6]. The purpose of these honeypots is to provide less risk in the network in the event of a compromise, but they provide little information about the attacker or the attack. In addition, production honeypots can be easily maintained and can provide security mechanisms such as detection, prevention, and reaction that can be used to protect the network of an organization. In contrast to a prevention system, research honeypots are primarily designed to gather information about the attackers, rather than to act as a deterrent. As an added benefit, this information can be helpful in analyzing the attackers, such as who the threat actors are, what their goals are, what tools or methods were used to attack a system, and so on. This type of information indirectly improves resource security in a number of ways. To better understand an attacker and their motives, we need something that resembles a real computer and operating system. Therefore, configuring research honeypots has become increasingly complex and time-consuming. As they are usually set up as “real” fake systems, this poses a greater risk to an organization and “potentially reduces security since they require extensive resources and maintenance” [5]. The level of interaction can also be used to categorize honeypots. There are typically three types of interaction: low interaction, medium interaction, and high interaction. Honeypots with low interactions usually lack an operating system. Thus, an attacker is therefore limited to only attempting logins. Among the services that may be enabled are Telnet, SSH, and FTP, but the honeypot doesn’t offer much else. Honeypots like this are typically produced in large production as they are easy to set up and are unlikely to be completely compromised by the attackers. Honeypots with medium interaction also tend not to have operating systems, similarly to low-interaction honeypots. While this may be true, they still tend to provide higher-level simulated services which are a desirable target for hackers. Moreover, these honeypots have a more complex structure, which makes them more difficult to configure. The reason for this is that they need to collect more information and understand how the attacker is going to behave during an attack, such as how the system was compromised and how tools and techniques are being used. The most complex systems are honeypots with a high degree of interaction. The process of installing and maintaining them can be considerably more difficult since they provide a more open environment for attackers to exploit. The use of these methods will allow us to collect much more data over time about attack behavior, as well as methodology. In general, higher rewards are associated with higher risks. Due to the fact that these systems are real, they can compromise an organization’s network. In order for these systems to operate in a secure environment, generally they need to operate behind a firewall within a controlled environment. Even though the attacker will be able to access the honeypot, the firewall will prevent him from benefiting further from it. Normally, these honeypots are used by researchers for the purposes of research. There is another classification that can be ascertained as there is a difference between a physical honeypot and a virtual honeypot. Honeypots that are physical would be mounted on the network of an organization, whereas honeypots that are virtual would be installed on a virtualization host using VirtualBox or VMware and look at any network traffic that is sent through the network. The advantage of virtual honeypots is the fact that they can be used to house a number of honeypots in a single system.

3 IoT Honeypot - An Application for Monitoring Internet of Things Attacks

A recent study was conducted by Kaspersky, which deployed 50 honeypots around the world for more than a year to collect valuable data on IoT attacks. In order to increase their data collection and eventually gain a better understanding, their honeypots utilized a variety of interaction levels. Furthermore, they cycled through the IP addresses constantly so that the honeypots would not be marked as such, thus lowering the number of attacks and the amount of useful information that could be gathered. There are approximately 20k infected sessions every 15 minutes [7] on average. According to Kaspersky's statistics, their Telnet honeypots were attacked over 105 million times from over 276,000 unique IP addresses. It is evident from the above that the same IP address is used multiple times, suggesting that IoT devices are being targeted continuously to be infected and attacked. It was found that Mirai malware was the most common threat to these IoT devices. Due to the fact that the malware has been freely available for a considerable period of time, as well as the code being capable of stringing together botnets of varying levels of complexity, this has occurred. About half of the attacks carried out against IoT honeypots [7] were made using the backdoor variant named "Backdoor.Linux.Mirai.c". There is no doubt that SSH, Telnet, and web servers are among the most commonly used and available services in the field of Internet of Things, which makes them an attractive target for attackers. In addition to this, it is also important to remember that IoT devices typically use a plethora of computing architectures that are quite different from those used by traditional computers. It is because of this that attackers are more likely to launch their malware once they have access to the honeypot and they are not checking for which architecture they are using. Basically, it works because only one line can be executed correctly at a time. As a result of this, researchers can trace back the sources of the attack tools used by attackers, which allows them to study them much more effectively later on.

3.1 IoT Honeypot - challenges

IoT devices pose a challenge when it comes to building honeypots if researchers rely on traditional methods because IoT devices have certain characteristics that need to be addressed. To maximize the chances of the hacker finding and exploiting vulnerabilities, it is important that the honeypot remains anonymous while mimicking a real system in order to prevent being easily identified by attackers. Due to the distinctive nature of IoT devices, as well as the inability to fully understand an attacker's nature and activities, an effective honeypot will need a different approach.

4 Explore the available Internet of Things Honeypots

In this section, we provide a brief overview of honeypot studies that may be applicable to general IoT use cases. In the first step, we identify a few classic honeypots that may be pretty applicable to the use case of general Internet of Things devices. The next section presents research on IoT honeypots, complete with emulation of all devices.

4.1 Classic Honeypots Overview

There is still the need to distinguish between honeypots that are designed for general purpose applications and those that are designed specifically for IoT applications. IoT honeypots inherit some characteristics from general application honeypots, including the ability to react to events as they arise. Despite the fact that these honeypots are not specifically designed for IoT, yet they are currently being used as research for IoT honeypots.

HoneyD [8] is an open-source program that is used to build scalable honeypots with low levels of interaction. Honeyd not only allows you to create virtual honeypots, but it also allows you to integrate machines as well. There are several protocols supported by this honeypot, including: UDP, TCP, FTP, SMTP, Telnet, IIS, POP, and telnet. Various studies have explored whether HoneyD can be used to create effective honeypots that attract attackers. Using IoT devices to simulate honeypots, the researchers compared them with real IoT devices in their study. Despite the similarity between the content served by honeypots and real devices, the average response time for queries and Nmap scans differed greatly.

Table 1: List of General Internet of Thing Honeypots

#	Honeypot Name	Level of Interaction	Simulated Services
1	HoneyD	Low	FTP, SMTP, Telnet
2	Dionaea	Medium	FTP, HTTP, MQTT, etc.
3	Adaptive Honeypot Alternative	Low/High	SSH
4	Cowrie	Medium	SSH

Dionaea is an open-source software that allows users to create medium interaction honeypots that simulate a variety of different services (such as FTP, HTTP, MQTT, etc). This application targets adversaries who attack hosts on the Internet with vulnerable services. Since adversaries attempt to install malware inside the Dionaea tool, the tool provides researchers with the ability to analyze malware and obtain a copy of that malware. [9]

Adaptive Honeypot Alternative In the paper titled "Self-Adaptive Honeypots Coercing and Assessing Attacker Behaviour", Wagener used both a low-interaction honeypot and a high-interaction honeypot to gather data. In using this data, he developed an SSH honeypot that depends on game theory and Machine Learning (ML) techniques, and it is called the Adaptive Honeypot Alternative (AHA). It is important to note that even though Wagener does not implement his honeypot in an IoT environment, it is used by other researchers as a foundation. As a result of their findings, the researchers discovered that attackers responded to adaptive honeypots by carrying out three times more interactions when responding to the customizable tools of the honeypot. This demonstrates the importance of adaptive honeypots in honeypot research. [10]

With *Cowrie*, you can create scalable, medium-to-high-interaction, virtual honeypots that can control and monitor a variety of behaviors. As a medium interaction honeypot, it logs an attacker's shell interaction on a simulated UNIX system via emulating several commands. As a high interaction honeypot, it is a proxy for SSH and Telnet to observe an attacker's interaction on another system. As a matter of fact, it acts as a proxy between an attacker and a pool of virtual machines that are configured in a backend server that allows flexible configuration. The Cowrie honeypot was forked off of the Kippo honeypot and simulates SSH, Telnet, SFTP, etc. services. [11]

Table 1 gives a list of some of the general honeypots for IoT that are considered.

4.2 Research Honeypots designed for the Internet of things

The most versatile IoT honeypots are capable of emulating any device that is connected to the internet. With full device emulation, it is more difficult for attackers to detect the honeypot, which adds greater realism to the honeypot. As part of this section, we only include honeypots that are capable of fully simulating any device. There are also IoT Honeypots that provide full device emulation. Honeypots for Internet of Things devices are presented in Table **??**, which perform a full emulation of IoT devices.

ThingPot [12] With the ThingPot platform, a complete IoT platform can be emulated and supported at an application level, ensuring that your IoT system is scalable, virtual, open-source, and scalable. Thingspot has been tested for 45 days using Extensible Messaging Protocol (XMPP) and a REST API. The majority of requests that were captured were HTTP REST requests. Researchers looked at the Internet for devices such as Philips Hue, Belkin, Wemo, and TPlink, and noted that the attackers were scanning for specific devices to attack, such as brute force attacks or fuzzing to gain control. To remain anonymous, the attackers also used a network called Onion Router (TOR) to achieve their goal.

IoTCandyJar [13] Luo et. al.'s honeypots can replicate IoT device behaviors without the risk of being compromised since they are intelligent and mimic the behavior of authentic IoT devices. They are referred to as intelligent interaction honeypots. Honeypots learn how to extend the session with attackers using Machine Learning with Markov Decision Processes and constantly analyze the behavior of IoT devices that are publicly accessible on the Internet, learning how best to extend the session with attackers. The IoTCandyjar platform collected 18 million raw requests during the study period, including about 1 million IoT requests. In terms of ports scanned, the most requests were for 80, 7547, 8443, 81, 8080, and 88. HTTP was the most commonly used protocol.

Multi-phased Multi-faceted IoT Honeypot Ecosystem [14] A study published in the ACM Conference on Computer and Communications Security(CCS), Tabari et. al. presented an approach to build a multi-phased and multi-faceted honeypot ecosystem where researchers observe the behavior of real-world attackers and gradually enhance the sophistication of a low-interaction IoT honeypot. The researchers

Table 2: List of Designed Internet of Thing Honeypots

#	Honeypot Name	Level of Interaction	Simulated Devices
1	ThingPot	Medium	Philips Hue, Belkin, Wemo, Tplink
2	IoTCandyJar	Low/Medium	General IoT devices
3	MP/MF Ecosystem	Low/Medium	D-link Camera, General IoT Devices
4	Honware	High	CPE devices

also developed a honeypot for IoT cameras that allowed them to gain an understanding of what attackers were trying to find in IoT cameras and get a better understanding of how they could defeat them. The "Honeycamera" is what they called it. The team created a proxy instance, called "ProxyPot", which was to be positioned between the IoT devices and the external networks and to facilitate analysts' study of the communication between the IoT devices and the external network. The preliminary results showed that the attacks became more sophisticated with each successive phase of the development process. Aside from this, they also captured activities that appear to involve direct human interaction rather than simply automated scripts.

Honware [15] Vetterl and Clayton, in their paper, presented a method for processing software images and extracting the file systems of these images, that presented a self-adapting platform that can emulate a wide range of IoT and customer premises equipment (CPEs). For the purpose of fully emulating devices, Honware uses a program called Quick Emulator (QEMU), which runs on a host operating system with a custom kernel and filesystem pre-built.

Table 2 gives a list of some of the research honeypots that have been specifically designed for IoT devices.

4.3 More IoT Honeypots

We briefly introduce some of the other honeypots that are available for IoT and CPS devices.

Conpot [16] is one of the most popular ICS honeypots that has been used by researchers over the years. The honeypot is an open-source, low-interaction honeypot that was developed under the Honeynet Project [17] and is being maintained to this day. There are many industrial protocols supported by Conpot, including Building Automation and Control Network, Guardian AST, Kamstrup, Modbus, S7comm, and many others like HTTP, FTP, SNMP, Intelligent Platform Management Interface, and TFTP. There are templates for Siemens S7 class PLCs, Guardian AST tank monitoring systems, as well as Kamstrup 382 smart meters provided in the package.

DiPot [18] A distributed ICS honeypot called DiPot was proposed by Cao et al. Based on Conpot honeypot framework, DiPot was developed. The framework is enhanced by adding improvements and enhancements to the Conpot framework in order to provide higher-fidelity simulations of ICS protocols, data collection and analysis as well as visualization and statistics support. According to the authors, DiPot honeypots have been found to successfully deceive Shodan search engines and be recognized as legitimate ICS devices around the world by DiPot honeypots installed within virtual machines in the cloud.

It is Ferretti and colleagues' objective [19] to analyze the Internet scanning traffic that is aiming at ICS in order to find out. Several low interaction Conpot honeypots were analyzed by the authors in order to study the scanning behavior of the scanners. It was configured in each honeypot to simulate a specific ICS device through a specific communication protocol (such as S7comm, Modbus/TCP, IEC-61850-104, EtherNet/IP, BACnet, HTTP, FTP, and SSH). According to their analysis, which covered a period of four months during which they operated the honeypots, most of the scanners were legitimate (e.g., Shodan, Censys, etc.) and showed certain patterns of scanning. In the authors' view, the use of legitimate scanner patterns could be used as a clue in detecting malicious scanning and attack activities directly targeting ICS environments.

XPOT [20] according to Lau et al., the XPOT mechanism has been designed as a honeypot that uses medium interactions for ICS. The XPOT device simulates Siemens S7-300 series PLCs and enables the attacker to compile, interpret, and load PLC programs onto the device. S7comm and SNMP protocols are supported by this honeypot.

ICS Honeypot using Honeyd [21] Disso et al. looked at SCADA security from the perspective of honeypots. Honeypots made of a real PLC device were tested as high-interaction honeypots, and honeypots built on Honeyd were tested as low-interaction honeypots. The Roo honeywall from the Honeynet

Project was placed in front of the honeypots. They measured the latency, the network traffic counters, and the background traffic levels (i.e., anti-honeypot techniques) to compare the high and low interaction honeypots.

GridPot [22] for SCADA systems, Redwood and his colleagues proposed a symbolic honeynet framework, namely SCyPH. As part of the proposed framework, a physics simulation of SCADA components will be incorporated with simulated SCADA system components and an anomaly detection system will be employed to detect deviations in the simulated data as a function of simulated simulations. To demonstrate GridPot, the authors used GridLab-D simulator for electric substation simulations and IEC 61850-based communication, as well as Newton-Raphson power flow solver algorithm for the voltage flow and current flow between the actors. Conpot was used to emulate IEDs, as well as the GOOSE/MMS and Modbus communication protocols for the communication between the devices.

5 Conclusion

Due to the rapidly growing adoption of IoT technology in society, the rates of attacks against IoT devices are increasing as well. As a result of the relatively new nature of the technology and the rush-to-market approach that most companies have to meet consumer demand, there are quite a few vulnerabilities that attackers can take advantage of. In the future, more research is necessary in order to construct more secure devices. It is possible for researchers to use honeypots to gain insight into and gain more knowledge about who and what the attackers are as well as how they gain access to these devices. It is critical that IoT users change their default passwords to something that is more complex and harder to guess, regularly update their firmware, reboot any device that is acting strangely, and restrict all of their IoT devices to a local VPN to prevent them from being exposed to the Internet.

References

[1] "How big is iot? 20.6 billion connected devices in 2020," Apr 2021.

[2] "Gartner says 5.8 billion enterprise and automotive iot endpoints will be in use in 2020."

[3] Contributor, "IoT healthcare to reach $136bn market value by 2021," Mar. 2016.

[4] A. Ziaie Tabari and X. Ou, "A first step towards understanding real-world attacks on iot devices," *arXiv preprint arXiv:2003.01218*, 2020.

[5] L. Spitzner, *Honeypots: Tracking Hackers*. USA: Addison-Wesley Longman Publishing Co., Inc., 2002.

[6] C. Seifert, I. Welch, P. Komisarczuk, *et al.*, "Honeyc-the low-interaction client honeypot," *Proceedings of the 2007 NZCSRCS, Waikato University, Hamilton, New Zealand*, vol. 6, 2007.

[7] S. Y. Dan Demeter and M. PREUSS, "Iot: A malware story."

[8] N. Provos, "Honeyd-a virtual honeypot daemon," in *10th DFN-CERT Workshop, Hamburg, Germany*, vol. 2, p. 4, 2003.

[9] DinoTools, "Dinotools/dionaea: Home of the dionaea honeypot."

[10] G. Wagener, "Self-adaptive honeypots coercing and assessing attacker behaviour."

[11] Cowrie, "Cowrie ssh/telnet honeypot https://cowrie.readthedocs.io."

[12] M. Wang, J. Santillan, and F. Kuipers, "Thingpot: an interactive internet-of-things honeypot," *arXiv preprint arXiv:1807.04114*, 2018.

[13] T. Luo, Z. Xu, X. Jin, Y. Jia, and X. Ouyang, "Iotcandyjar: Towards an intelligent-interaction honeypot for iot devices," *Black Hat*, pp. 1–11, 2017.

[14] A. Ziaie Tabari and X. Ou, "A multi-phased multi-faceted iot honeypot ecosystem," in *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*, CCS '20, (New York, NY, USA), p. 2121–2123, Association for Computing Machinery, 2020.

[15] A. Vetterl and R. Clayton, "Honware: A virtual honeypot framework for capturing cpe and iot zero days," in *2019 APWG Symposium on Electronic Crime Research (eCrime)*, pp. 1–13, IEEE, 2019.

[16] Conpot.

[17] L. Spitzner, "The honeynet project: Trapping the hackers," *IEEE Security & Privacy*, vol. 1, no. 2, pp. 15–23, 2003.

[18] J. Cao, W. Li, J. Li, and B. Li, "Dipot: A distributed industrial honeypot system," in *International Conference on Smart Computing and Communication*, pp. 300–309, Springer, 2017.

[19] P. Ferretti, M. Pogliani, and S. Zanero, "Characterizing background noise in ics traffic through a set of low interaction honeypots," in *Proceedings of the ACM Workshop on Cyber-Physical Systems Security & Privacy*, pp. 51–61, 2019.

[20] S. Lau, J. Klick, S. Arndt, and V. Roth, "Poster: Towards highly interactive honeypots for industrial control systems," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, pp. 1823–1825, 2016.

[21] J. P. Disso, K. Jones, and S. Bailey, "A plausible solution to scada security honeypot systems," in *2013 Eighth International Conference on Broadband and Wireless Computing, Communication and Applications*, pp. 443–448, IEEE, 2013.

[22] O. Redwood, J. Lawrence, and M. Burmester, "A symbolic honeynet framework for scada system threat intelligence," in *International Conference on Critical Infrastructure Protection*, pp. 103–118, Springer, 2015.