Design Characteristics of Honeypots - with an Aim to Study Botnets

by

Amber Higgins, B.A., B.A.I.

A Dissertation submitted to the University of Dublin, Trinity College

in fulfillment of the requirements for the degree of

Integrated Masters in Computer Engineering

May 2018

Declaration

I, the undersigned, declare that this work has not previously been submitted as an exercise for a degree at this, or any other University, and that unless otherwise stated, is my own work.

Amber Higgins

March 11, 2018

Permission to Lend and/or Copy

I, the undersigned, agree that Trinity College Library may lend or copy this thesis upon request.

Amber Higgins

March 11, 2018

Acknowledgments

This research has been a significant undertaking that I could not have accomplished without the support and guidance of many individuals.

First and foremost my thanks must go to my supervisor Dr. Stefan Weber, who has consistently been patient and supportive of me through all the issues that I have encountered. His contribution has been absolutely exceptional.

Special mentions go to Jason Flood of IBM Ireland, Michel Oosterhof of the Cowrie project, and Tony Winters of Optum Ireland who offered valuable insights and guidance on the way forward at various points in this research.

Lastly and most importantly of all, to my loving family - Mom, Dad, Romy and Holly - and my partner Breandan. I could never accomplish any of this without you.

Amber Higgins

University of Dublin, Trinity College

May 2018

Abstract

The abstract or summary or whatever...

Contents

Acknowledgments **iv**

Abstract **v**

List of Tables **ix**

List of Figures **x**

Chapter 1 Introduction **1**

1.1 Problem Area 1

1.2 Research Objectives 2

1.3 Contributions 3

1.4 Report Structure and Contents 3

Chapter 2 State of the Art **4**

2.1 Cyber Threats in 2018 4

2.1.1 Evolution of the Cyber-Threat Landscape 4

2.1.2 Motivations 5

2.1.3 The Internet of Things 5

2.2 Botnets 5

2.2.1 Background 6

2.2.2 Design 7

2.2.3 State of the Art 7

2.3 Honeypot Technologies . . . 8
2.3.1 Background . . . 8
2.3.2 Design . . . 9
2.3.3 State of the Art . . . 11
2.4 Cyber Incident Monitoring . . . 13
2.4.1 Definition . . . 13
2.4.2 Motivations . . . 13
2.4.3 Challenges . . . 14
2.4.4 The Role of Honeypots . . . 14
2.5 Closely-related Work . . . 14
2.5.1 IoTPot . . . 14
2.5.2 IoTCandyJar . . . 15
2.5.3 WSN Honeypot . . . 15
2.5.4 Distributed Virtual Honeynets . . . 15
2.5.5 Honeypot-Based Cyber-Incident Monitors . . . 16
2.6 Summary . . . 16

Chapter 3 Design **17**

3.1 Proposed Work . . . 17
3.1.1 Cyber-Incident Monitor for Critical Infrastructures . . . 17
3.1.2 Tracking Internet-of-Things Botnets . . . 17
3.2 Design Decisions . . . 17
3.2.1 Considerations . . . 17
3.2.2 Challenges . . . 17
3.3 Summary . . . 17

Chapter 4 Implementation **18**

4.1 Overview . . . 18
4.2 Detail 1 . . . 18

4.3 Detail 2 18
4.4 Summary 18

Chapter 5 Evaluation **19**

5.1 Experimental Setup 19
5.2 Experiment 1 19
5.3 Experiment 2 20
5.4 Experiment 3 20
5.5 Discussion 20
5.6 Summary 20

Chapter 6 Conclusions & Future Work **21**

6.1 Conclusions 21
6.2 Future Work 21
6.2.1 Extension of the Cowrie Project 21
6.2.2 Alternative Network Configurations 22
6.2.3 Extension of Log Analysis 22

Appendix A Abbreviations **23**

List of Tables

2.1 ToC Caption 16

5.1 Variables of the experiment 19

List of Figures

1.1 An Image of a chick 2

5.1 Measurement of System Wakeups 20

Chapter 1

Introduction

The section following the chapter title should give an overview of overview of the chapter in 1 to 3 paragraphs. In any document, a title should always be followed by text; a title should never be followed immediately by another title.

1.1 Problem Area

Botnets are continuing to make headlines. They are being used in more diverse ways than ever before.

Critical service infrastructures are being widely targeted. It is becoming a highly profitable business for criminals to rent out botnets to political organisations and others (reference booter services for DDoS - a form of modern-day protest?) There are more internet connected devices than ever before more devices to be exploited, compromised and harnessed for their computational resources. As aptly described by leading cybersecurity expert Bruce Schneier in his article Botnet of Things [2], botnets will get larger and more powerful simply because the number of vulnerable devices will go up by orders of magnitude over the next few years ... overall, the trends favor the attacker.

These huge numbers of internet-connected devices are increasingly being operated by individuals with very limited technical knowledge, and even less awareness of security.

These huge numbers of interconnected devices with poorly implemented security measures are an ideal playground for hackers.

Honeypots allow the would-be victims of these cybercriminals to learn first-hand about the techniques and motivations of the attackers. The importance of this kind of active defence mechanism is growing as attacks are becoming more and more unpredictable.

An example of how to reference a figure in the thesis document; see figure 1.1.

Figure 1.1: This caption should describe the figure to the reader and explain to the reader the meaning of the figure. If the interpretation of a figure is left to the reader, the reader will misinterpret the figure!

1.2 Research Objectives

This project is being undertaken in a bid to understand the role of honeypot technology in emerging cybersecurity use-cases in particular, in tracking and defending against botnets.

The primary objectives of the research are to:

1. Develop a novel, honeypot-based cyber-incident monitoring system to provide effective intrusion detection and attack mitigation for critical infrastructures; and

2. Obtain a deeper understanding of emerging Internet-of- Things (IoT) botnets using targeted honeypots deployments, enhancing the knowledge of the wider cybersecurity community.

With the evolution of new threats to the security of data and systems every day as technology continues to advance, these objectives are conducive to the longer-term fight against cyber- exploitation in general. Proposing a novel and effective system to address the two objectives is the central aim of this research.

1.3 Contributions

The contributions of the work to the field of knowledge.

1.4 Report Structure and Contents

A description of the structure of the dissertation that explains to the reader the contents of the following chapters and the thoughts behind the layout of the dissertation. The chapters following the introduction...

Chapter 2

State of the Art

This chapter should explain the existing work of the areas that your work is based on. The introduction should explain to the reader the area of the work as a whole, how the individual area contribute to the work and what the reader will find in the discussion of each of the areas. The idea is that the reader will be aware of the general contents of the state of the art and will not be surprised by any of the issues that are being discussed.

2.1 Cyber Threats in 2018

Explanation of existing work of a given area that describes the area as a whole, how it fits into the work and then breaking it down into components that are relevant to the work.

An example for possible citations (**?**) or by **?**.

2.1.1 Evolution of the Cyber-Threat Landscape

Starting with a general description of the issue; then drilling down into the details of the issue and how it has been covered in the literature. For a skeleton at the beginning of the writing, this text should be replaced by a general short description, so that you know what you want to discuss and can review the sequence of the discussion.

2.1.2 Motivations

Starting with a general description of the issue; then drilling down into the details of the issue and how it has been covered in the literature. For a skeleton at the beginning of the writing, this text should be replaced by a general short description, so that you know what you want to discuss and can review the sequence of the discussion.

2.1.3 The Internet of Things

Connectivity has been introduced rapidly into processes in critical systems and infrastructures, including power, agriculture, health and transport. Now, devices and processes which were never online before are now being exposed to external attacks, allowing them to be significantly tampered with in ways that can affect the welfare of the entire population that relies on them. Crucial systems and infrastructures on which society relies on a daily basis such as satellites, power grids, medical devices and traffic control systems are connected to the internet, making them all vulnerable to potential attacks.

IoT devices will continue to be one of the most effective and easily exploitable tools for attackers. Until there is a truly commercial, market-driven need to secure IoT devices, manufacturers will continue to avoid the expense of recalling and securing the billions of IoT devices that have been so far exploited by these attackers. This means that unfortunately, the exploitation of IoT devices will not end any time soon.

2.2 Botnets

Explanation of existing work of a given area that describes the area as a whole, how it fits into the work and then breaking it down into components that are relevant to the work.

An example for possible citations (**?**) or by **?**.

2.2.1 Background

Starting with a general description of the issue; then drilling down into the details of the issue and how it has been covered in the literature. For a skeleton at the beginning of the writing, this text should be replaced by a general short description, so that you know what you want to discuss and can review the sequence of the discussion.

Definition

Motivations

Generally, botnets are used for any tasks that require volume, whether this is volume of network traffic, computation, or something else.

•$Distributed Denial of Service (DDoS)$

A number of machines under the control of the botnet are used to simultaneously overwhelm a website or server with traffic, resulting in performance degradation and even failure of the service.

•$Click fraud$

Botnets can be used to generate artificial user interactions with advertisements on the web, making it appear that real users have clicked on an advertisement for the purpose of generating revenue.

•$Extortion$

IP cameras being hacked: make reference to the website

•$Spambots$

•$Password - cracking$

Since botnets allow for synchronised computations across multiple devices, they are often used for brute-force password-cracking computations.

•$Bitcoin mining$

The explosion of interest in cryptocurrencies over the last several months has led to a significant increase in the number of botnets being used to mine bitcoin and other cryp-

tocurrencies. Botnets are also being used for DDoSing of competing groups performing mining.

Make reference to cryptojacking trends!

2.2.2 Design

Starting with a general description of the issue; then drilling down into the details of the issue and how it has been covered in the literature. For a skeleton at the beginning of the writing, this text should be replaced by a general short description, so that you know what you want to discuss and can review the sequence of the discussion

Architecture

Centralised

Peer-to-Peer

2.2.3 State of the Art

Starting with a general description of the issue; then drilling down into the details of the issue and how it has been covered in the literature. For a skeleton at the beginning of the writing, this text should be replaced by a general short description, so that you know what you want to discuss and can review the sequence of the discussion

Recent Developments

Emerging Trends

Timeline of Evolution

2.3 Honeypot Technologies

2.3.1 Background

Definitions

Cybersecurity expert and Honeynet Project founder Lance Spitzner defines honeypots as quot;a security resource who39;s value lies in being probed, attacked or compromisedquot; [3]. They are categorised as an active network defence mechanism, and are deployed solely with the intention that they will be attacked.

Honeypots can be classified as both decoys and sensors depending on the context of their deployment.

•*Decoys*

A honeypot will divert an attackers attention away from valuable devices in the network, particularly important in a production environment. In order for this strategy to be effective, the honeypot should be the perfect decoy, appearing to be exactly what the attacker is looking for.

•*Sensors*

Honeypots always act as sensors in the network in which they are deployed. When attacked, system administrators can be alerted and valuable attack data can also be captured for analysis regarding the attackers behaviour, strategies and motivations.

Honeynets TODO define and explain.

Motivations for Use

Traditional intrusion-detection mechanisms, such as firewalls, define all attackers passively and simply alert to the fact that a security incident, such as an unauthorised connection attempt, has occurred. Honeypots differ in this regard: They proactively entice attackers to give away their attack strategies and intentions by leaving their tracks behind. By learning what these attackers are targeting, more effective defences can be implemented as vulnerabilities and flaws are identified.

There are a number of specific and significant benefits associated with the use of honeypot technologies.

•$Mimicexistingsystems$

•$Insidethenetwork$

•$ExtensiveLoggingPotential$

•$Goal-oriented$

•$FewFalsePositives$

A premise of using a honeypot is that nobody should communicate with it: There is no legitimate reason to interact with a honeypot. This means that typically, a honeypot will trigger very few false positives, since every interaction with a honeypot is automatically distrusted. It is important however to realise that in a production environment, individuals within the organisation may mistakenly interact with the honeypot something which can often be deduced from looking at the logging captured by higher-interaction honeypots.

•$OffensiveandDefensive$

2.3.2 Design

Interactivity

Interactivity levels of honeypots are an important consideration, which define (i) the ability of the attacker to interact with the honeypot, and consequently (ii) the volume

and type of information that can be gathered by that honeypot. Levels of interaction range from simply allowing a connection to be made over SSH or another protocol, to being able to download and install malware binaries. The cost versus learning benefits of honeypot interactivity levels increase proportionally, meaning that a highly interactive honeypot will likely be expensive to host and maintain.

There are three typical classifications of honeypot interactivity level: Low, high and medium.

- Low

  These honeypots are most commonly used for intrusion detection, i.e. they are designed to alert someone to the fact that an attack has occurred, but do not provide any means of interacting with the attacker or capturing the attack data.

- High

- Medium

  The required level of interactivity of a honeypot with their attacker depends on the use context: Low-interaction devices are commonly used for intrusion detection, whereas high-interaction devices are live systems that can capture detailed information about the behaviour of attackers such as botnets in the wild.

Deployment Scenarios

- Production Honeypots

  Production honeypots are generally deployed in large, corporate networks with the intention that they will act as a part of the active network defence in the organisations infrastructure. Their primary purpose is to act as a decoy, luring the attacker away from valuable machines on the network with a seemingly more valuable and vulnerable target on the network. This enables the honeypot to alert system administrators to the fact that there has been an intrusion, giving them the chance

to isolate the valuable devices in the network from the infected honeypot. It also enables system administrators to identify vulnerable points in their infrastructure, allowing them to continuously improve the security of their systems.

The idea of attracting attackers into the system, making them interact with it and give away their attack strategies without causing any harm to the real systems is one of the major attractions of using honeypots in a production environment. The solution being proposed as part of this research is targeted at use in production environments for exactly this reason, thwarting the attackers intended compromise of the network.

- Research Honeypots

  As the name suggests, these honeypots are generally deployed for the purpose of research rather than as a security measure to protect a network. The emphasis with these honeypots is not so much on the ability of the honeypot to act as a decoy node in a network, and more on the ability of the honeypot to collect valuable data for analysis. By allowing attackers to interact with and infect the honeypot, research can be carried out relating to the behaviour and strategies of the attackers.

Applications

2.3.3 State of the Art

Honeypots of Note

At the time of writing, there were over X honeypot projects listed.

1. Cowrie

   Cowrie is a medium interaction honeypot which was originally forked from the Kippo honeypot project. It is under active development and is widely used as an SSH/Telnet honeypot solution, providing all of the features that the Kippo honeypot provided and more.

It has many useful capabilities and features, including the following:

- It is actively maintained by a dedicated community, with quick response times to queries and bug-fixes.
- It enables attackers to gain access to the honeypot using SSH and Telnet, both protocols widely exploited by IoT botnets.
- Cowrie fully emulates a Debian installsation, with an out-of-the-box configurable filesystem that an attacker can interact with.
- It is an open-source project, allowing users to adapt the honeypot to suit their specific needs.
- Cowrie records interactions between the attacker and the honeypot, logging everything from commands executed and file downloads attempted to the source IP addresses and protocol information.
- Cowrie does not limit the number of simultaneous sessions, and can present a mock filesystem and shell to each attacker independently.
- The fact that the Cowrie software doesn't call on any external sofware to operate makes it much less vulnerable to third-party compromises. It also improves substantially on its predecessor, Kippo, in that many of the fingerprinting issues are resolved.

  Cowrie cannot execute malware, but can be used in conjunction with other solutions such as Cuckoo Sandbox in order to execute malware safely in a controlled environment.

2. Dionaea

Dionaea is a commonly used low-interaction honeypot. It is a medium-interaction honeypot, which emulates a vulnerable Windows operating system running protocols like SSH, HTTP and FTP. The stated goal of Dionaea is to trap malware and obtain a copy of it, to allow it to be inspected using other software such as Cuckoo Sandbox.

3. Kippo

   Kippo is a honeypot written in Python which is no longer under active development, but whose direct descendants include the Cowrie honeypot. It is a medium-interaction SSH honeypot, which is designed to capture the entire session interaction of a connected attacker. It is an emulation of a Debian Linux installation, which provides a fake file system and shell for each attacker to interact with while logging all attack data including records of any downloaded malware.

   Kippo includes configuration options for permitted usernames and password combinations, and all brute-force attempts are logged. As quoted from desaster, creater of Kippo, "By running kippo, you're virtually mooning the attackers. Just like in real life, doing something like that, you better know really well how to defend yourself!"

4. Honeyd

Challenges

2.4 Cyber Incident Monitoring

Explanation of existing work of a given area that describes the area as a whole, how it fits into the work and then breaking it down into components that are relevant to the work.

An example for possible citations (**?**) or by **?**.

2.4.1 Definition

Starting with a general description of the issue; then drilling down into the details of the issue and how it has been covered in the literature. For a skeleton at the beginning of the writing, this text should be replaced by a general short description, so that you know what you want to discuss and can review the sequence of the discussion.

2.4.2 Motivations

Starting with a general description of the issue; then drilling down into the details of the issue and how it has been covered in the literature. For a skeleton at the beginning of the writing, this text should be replaced by a general short description, so that you know what you want to discuss and can review the sequence of the discussion.

2.4.3 Challenges

Starting with a general description of the issue; then drilling down into the details of the issue and how it has been covered in the literature. For a skeleton at the beginning of the writing, this text should be replaced by a general short description, so that you know what you want to discuss and can review the sequence of the discussion.

2.4.4 The Role of Honeypots

Starting with a general description of the issue; then drilling down into the details of the issue and how it has been covered in the literature. For a skeleton at the beginning of the writing, this text should be replaced by a general short description, so that you know what you want to discuss and can review the sequence of the discussion.

2.5 Closely-related Work

There is a renewed focus on harnessing the unique capabilities of honeypot technologies in light of the volume of recent and severe cyber-attacks. Honeypots have even been commercialised in some instances [4] [5], such has been the consideration of their role in intrusion detection in production environments. However as Canary highlights on their product landing page, using honeypots in large networks doesnt happen practically because with all the network problems we have, nobody needs one more machine to administer and worry about. This highlights clearly both the demand and lack of supply

for a solution for large organisations regarding practical deployment and maintenance of honeypot-based frameworks.

2.5.1 IoTPot

Studying IoT botnets using targeted honeypot deployments has been investigated by several researchers. Yin Minn Pa Pa et al. propose IoTPot [6], a reactive, adaptable IoT honeypot that responds to architecture-specific requests with corresponding architecture-specific responses.

2.5.2 IoTCandyJar

Another IoT honeypot, IoTCandyJar [7] aims to be an intelligent-interaction honeypot, capturing attack data more effectively by utilising machine learning techniques to determine the most appropriate behaviour to prolong attacks.

However, for both of these solutions the reactivity of the honeypot to probing and attack is the only characteristic studied for effectiveness.

2.5.3 WSN Honeypot

In contrast, S. Dowling et al. propose Zigbee Honeypot [8] to assess IoT cyber- ttacks launched on wireless sensor networks (WSNs). The focus of this implementation is on attracting these attacks, rather than identifying the characteristics of the Zigbee honeypot that make it most susceptible to attack: There are however a number of useful tactics used to attract the desired attackers. In relation to honeynets, there has been substantially less work done on developing a system suitable for large-scale deployment.

2.5.4 Distributed Virtual Honeynets

Pisarck et al. [9] investigate the effectiveness of a distributed virtual honeynet, using OS-level virtualisation. However, their investigation is limited to simulations of botnet

DDoS attacks, limiting the applicability of their system to a live production environment.

Kedrowitsch et al. [10] also look into the benefits and drawbacks of using OS-level virtualisation in honeypot deployments, providing many useful insights on the viability of this approach.

2.5.5 Honeypot-Based Cyber-Incident Monitors

A honeypot-based cyber-incident monitor is proposed by Vasilomanolakis et al., which recognises the importance of data aggregation and visualisation from multiple honeypots in order to provide important data to key decision makers. [11] This cyber-incident monitor is not targeted at the use- case of critical service infrastructures or low-cost intrusion detection solutions, but addresses the major consideration in this project of centralising important data for key decision-makers.

Deutsche Telekom (DT) have also developed an open-source honeypot-based visualisation tool called T-Pot [12] using dockerised honeypots to allow organisations to manage honeypot deployments more easily. DT use this framework to manage the security of their own telecommunications infrastructure, and their work on this project strengthens the case for the coupling of scalable honeynet deployments and visualisation tools.

2.6 Summary

This section should summarize the essential points of the state-of-the-art and give the reader an overview of relevant projects; ideally, this can be summed up by providing a table (see table 2.1) with the relevant projects and issues that they address.

Column 1	**Column 2**	**Column 3**	**Column 4**
Row 1	Item 1	Item 2	Item 3
Row 2	Item 1	Item 2	Item 3
Row 3	Item 1	Item 2	Item 3
Row 4	Item 1	Item 2	Item 3

Table 2.1: Caption that explains the table to the reader

Chapter 3

Design

Description of the Design chapter and its contents.

3.1 Proposed Work

General discussion of the design and the overall decisions that were made.

3.1.1 Cyber-Incident Monitor for Critical Infrastructures

3.1.2 Tracking Internet-of-Things Botnets

3.2 Design Decisions

3.2.1 Considerations

3.2.2 Challenges

3.3 Summary

Chapter 4

Implementation

4.1 Overview

4.2 Detail 1

```
x = 1
if x == 1:
    # indented four spaces
    print("x is 1.")
```

Listing 4.1: Lengthy caption explaining the code to the reader

4.3 Detail 2

4.4 Summary

Chapter 5

Evaluation

5.1 Experimental Setup

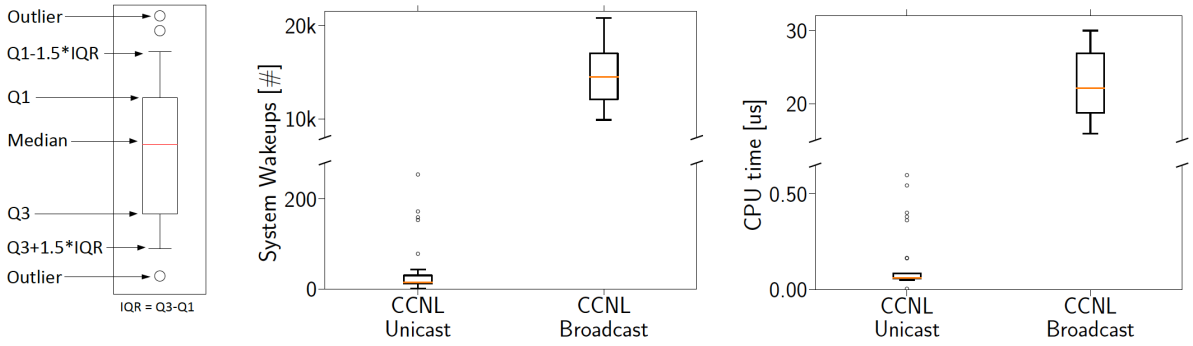
Describes the experimental setup and the values that were defined for the variables as given in table 5.1.

Column 1	Column 2
Row 1	Item 1
Row 2	Item 1
Row 3	Item 1
Row 4	Item 1

Table 5.1: Caption that explains the table to the reader

5.2 Experiment 1

Figure 5.1 shows measurements.



(b) System wakeups (c) CPU usage

Figure 5.1: Long caption that describes the figure to the reader

5.3 Experiment 2

5.4 Experiment 3

5.5 Discussion

5.6 Summary

Chapter 6

Conclusions & Future Work

6.1 Conclusions

6.2 Future Work

There is substantial scope for additional developments based on the foundations laid by this research.

6.2.1 Extension of the Cowrie Project

During the course of the research, it was discovered that the Cowrie honeypot did not provide the capability to make outgoing network connections. By design, medium-interaction honeypots should not give an attacker the ability to use all functionality that would exist in a real environment - however, this limitation restricted the scope of the work during the project.

Given the time within which the research was being undertaken, it was not possible to extend the Cowrie source code to give it the ability to make outbound network connections over SSH and Telnet. Potential changes to the source code were documented, and are included in the following sections.

Outbound SSH Connectivity

To be completed.

Outbound Telnet Connectivity

To be completed.

6.2.2 Alternative Network Configurations

6.2.3 Extension of Log Analysis

Appendix A

Abbreviations

Short Term	**Expanded Term**
DNS	Domain Name System
DHCP	Dynamic Host Configuration Protocol
...	...