Gaining Cyber Security Insight Through an Analysis of Open Source Intelligence Data: An East African Case Study

Submitted in fulfilment
of the requirements of the degree of

Master of Science

of Rhodes University

Stones Dalitso Chindipha

Grahamstown, South Africa
December 2016

Abstract

With each passing year the number of Internet users and connected devices grows, and this is particularly so in Africa. This growth brings with it an increase in the prevalence cyber-attacks. Looking at the current state of affairs, cybersecurity incidents are more likely to increase in African countries mainly due to the increased prevalence and affordability of broadband connectivity which is coupled with lack of online security awareness. The adoption of mobile banking has aggravated the situation making the continent more attractive to hackers who bank on the malpractices of users.

Using Open Source Intelligence (OSINT) data sources like Sentient Hyper-Optimised Data Access Network (SHODAN) and Internet Background Radiation (IBR), this research explores the prevalence of vulnerabilities and their accessibility to cyber threat actors. The research focuses on the East African Community (EAC) comprising of Tanzania, Kenya, Malawi, and Uganda. An IBR data set collected by a Rhodes University network telescope spanning over 72 months was used in this research, along with two snapshot period of data from the SHODAN project.

The findings shows that there is a significant risk to systems within the EAC, particularly using the SHODAN data. The MITRE CVSS threat scoring system was applied to this research using FREAK and Heartbleed as sample vulnerabilities identified in EAC. When looking at IBR, the research has shown that attackers can use either destination ports or IP source addresses to perform an attack which if not attended to may be reused yearly until later on move to the allocated IP address space once it starts making random probes. The moment it finds one vulnerable client on the network it spreads throughout like a worm. DDoS is one the attacks that can be generated from IBR.

Since the SHODAN dataset had two collection points, the study has shown the changes that have occurred in Malawi and Tanzania for a period of 14 months by using three variables i.e. device type, operating systems, and ports. The research has also identified vulnerable devices in all the four countries. Apart from that, the study identified operating systems, products, OpenSSL, ports and ISPs as some of the variables that can be used to identify vulnerabilities in systems. In the case of OpenSSL and products, this research went further by identifying the type of attack that can occur and its associated CVE-ID.

Acknowledgements

First and foremost I would like to thank God for the gift of life and ability work tirelessly while working on this thesis. I also would like to acknowledge and extend my gratitude to a couple of people who gave me support and guidance throughout the duration of this research.

This research would not have been possible without the constant guidance and support of my supervisor, Prof Barry Irwin. His constant desire for perfect work pushed me to the limit and guided me to the success and completion of this research. I would like to also extend thanks to my mother for her continuous support, motivation and believing in me that this work will be done and praying for me when I seemed to despair with me being miles away from home for two years.

Special thank to Alan Herbert, Akhona Ngqolongo and Nhlakanipho Dlamini who took time in their busy academic and work schedules to proofread this work and point out parts that needed further explanation due to either poor technical use of terms or grammatical issues which in the end made this more readable and complete. My lab mates also deserve special mention, Sean Pennefather and Lauren Rudman who both made the lab more conducive to work in but also offered assistance where I fell short.

I would like to thank the Computer Science Department at Rhodes University for providing me access to the equipment and space required to carry out this research. Finally, I wish to gratefully acknowledge the support of Rhodes University and the Beit Trust who coordinated for the financial support for this research.

Contents

List of Tables

List of Figures

Listings

Glossary

The following list describes the various abbreviations and acronyms used throughout this document. Citations of research related to these areas listed below can be found in the text to better retain the context upon which the citation relies.

API Application Programming Interface

APT Advance Persistent Threats

ARP Address Resolution Protocol

CCS Change Cipher Spec

CII Critical Information Infrastructure

CIA Confidentiality, Integrity and Availability

CIDR Classless Inter-Domain Routing

CSV Comma Separated Values

CSIRTs Computer Security Incident Response Teams

CVE Common Vulnerability Exposure

CVSS Common Vulnerability Scoring System

C&C Command and Control

DBT Design Based Threat

DCOM Distributed Component Object Model

DDoS Distributed Denial of Service

DIDS Distributed IDS

DoS Denial of Service

DTLS Datagram Transport Layer Security

EAC East African Community

FIRST Forum of Incident Response and Security Teams

FTP File Transfer Protocol

HTTP Hyper Text Transfer Protocol

HTTPS Secure Hyper Text Transfer Protocol

IANA Internet Assigned Number Authority

IBR Internet Background Radiation

ICMP Internet Control Message Protocol

IDS Intrusion Detection System

IP Internet Protocol

ISP Internet Service Provider

JSON JavaScript Object Notation

KB kilobytes

MAC Media Access Control

IIS Microsoft Internet Information Services

MB Megabytes

MitM Man in the Middle

MTL Malawi Telecommunications Ltd

OCSP On-line Certificate Status Protocol

OpenSSL Open Secure Sockets Layer

OSINT Open Source Intelligent

PBX private branch exchange

pcap packet capture

PKI Public Key Infrastructure

RDP Remote Desktop Protocol

RPC Remote Procedure Call

RSA Relatively Slow Algorithm

SHODAN Sentient Hyper-Optimized Data Access Network

SMS Short Message Service

SMTP Simple Mail Transfer Protocol

SSL Secure Sockets Layer

SQLI Structured Query Language injection

TCP Transmission Control Protocol

TLS Transport Layer Security

TTCL Tanzania Telecommunications Limited

UDP User Datagram Protocol

WAFWs Web Access Firewall

WAP Wireless Application Point

XSS cross site scripting

Chapter 1

Introduction

The 20th century saw many technological advancements through innovation. Among such innovations is the use of the Internet. The Internet has opened up new possibilities for accessing information and has had more influence than any form of intelligence preceding it; like cultural intelligence in problem solving. This has however brought in information security risks and privacy issues as it renders resources easily accessible to all people. Risks include breach of access, spamming, spyware which all aim at deliberately compromising networks and computer systems. That is not all there is to such publicly accessible information. Governments have used this publicly accessible information to make a strategic decision and launch wars against other nations. A good example is the case of Iran, where the Stuxnet worm exploited vulnerabilities in Iranian enrichment facilities affecting its nuclear potential (O'Mahony, 2015).

Being a public resource, Internet has been used by different groups of people to meet their goals. Companies, for example, use it to make business decisions like assess competitors on the market and perform due diligence on potential clients (Ponder-Sutton, 2016). However, it is not always the case that what Internet users are looking for can easily be found with just a Google search or looking at social media. There is a need for special tools that are needed to access certain information that may not otherwise be available. To make sense of large data acquired on the Internet there comes a need to use special tools to first access it, then process it for the intended purpose. This does not change the fact that the intelligence is publicly available to people, it just makes it relatively harder to access. Such tools for information gathering include creepy, TheHarvester[1],

[1] http://www.rwbnetsec.com/theharvester/

SHODAN, Search Diggity[2] and Recon-ng[3] all of which operate on different computer platforms (Tsalis and Gritzalis, 2015).

Cyber threat actors have been on the leading front in accessing such tools in order to conduct a reconnaissance of their targets however applying security intelligence into the tools protecting the organization is what is required by security professionals in various organisations with the aim of changing an approach: from one that is reactive to security attacks to one that proactive to them. Getting the right type of intelligence from the right sources and being able to effectively apply it to its associated environments is a critical component in the overall protection of the organizations and user of the Internet in general.

1.1 Problem Statement

The Internet is beneficial to our daily activities and with each passing day it also brings along information security concerns to the users, be it at a company or national level (Swart *et al.*, 2014). The rapid development of Internet access throughout Africa has not been accompanied by an equivalent increase in awareness of security issues, and that has opened up the possibility of a rise in cyber attacks. With each passing year the number of Internet users keeps growing in Africa (Chindipha and Irwin, 2015; Jensen, 2015). Undersea cables in Africa have risen from 6 active undersea cables as of 2010-2011 to 10 undersea cables currently resulting in gaining 14.020 GB/s (Van Niekerk and Maharaj, 2013). As of this writing, Africa's international bandwidth has reached 7 Terabytes Per Second (TBps)[4] meaning that there is more bandwidth for Internet users now than it was three years ago. Much as this is good news for development in Sub-Saharan region, it comes along with it a huge price of an explosion of cyber-crime activities and malware; which comes due to lack of awareness to information security issues.

Such Internet connectivity has made more resources easier to access, affordable and quicker to attain for everyone. Cyber threat actors like hackers often conduct reconnaissance when they opt to attack a potential target. It has been found that often such information is available to the public to an extent that every detail the attacker needs to access classified

[2]https://www.bishopfox.com/blog/2014/08/searchdiggity-avoid-bot-detection-issues-leveraging-google-bing-shodan-apis/

[3]https://bitbucket.org/LaNMaSteR53/recon-ng

[4]http://www.africabandwidthmaps.com/

information for the intended target is available to them and it is easy to create a profile for the target (Tsalis and Gritzalis, 2015).

It is hard to imagine now how hackers would operate without Open Source Intelligence (OSINT) as it has formed a fundamental precedence that cannot be overlooked any longer. OSINT is defined as any unclassified intelligence or information that is collected and acquired through the use of publicly available resources (Steele, 2007). Infosec institute confirmed on their website that an estimated 80% of critical information required to perform a cyber attack is available for use in OSINT. Specific information vital for a deep analysis attack is found in newspapers, Social network websites, magazines, web content, academic publications, industry newsletters, television transcripts, and blogs. All these make penetration testing which took hours for information gathering alone, to only takes a couple of minutes with OSINT. In short, we are more exposed to cyber threats because of OSINT than any other means of information gathering.

While this is the case, OSINT is an important component for understanding human problem solving in the 21st century. For many people, the Internet is limited to the results of a preferred search engine but there are other forms of search engines whose sole purpose is to scan the Internet and grab service banners based on IP address and port. One of such systems is SHODAN which automates the process of data extraction and exposes system weaknesses making reconnaissance trivial (Tsalis and Gritzalis, 2015).

Using SHODAN, this study will seek insight into identifying some of the variables that ought to be given special attention. This will be substantiated by the use of historical Internet Background Radiation (IBR) data acquired from network telescopes (Irwin, 2011).

1.2 Research Goals

This research aims to investigate and achieve the following goals:

1. Demonstrate using Common Vulnerability Scoring System (CVSS) how vulnerabilities can be quantified so as to provide a gauge of potential damage that can be caused by it and to show levels of significance that a given vulnerability needs given a specific CVE-ID identified from the SHODAN data found. Note that the CVE-ID has to be identified within the data set itself.

2. Using an appropriate IBR data set, identify variables that can be used to assess levels of vulnerability in the East African community. The study will use the identified variables to assess how each of the four countries is faring along the lines of security and identify those that are frequently being used.

3. Since data will be collected from these countries at different points, the study will conduct a comparative study analysis to observe if there has been a change during the study and assess if the change is for better or worse. This will be possible since the study uses Open Source Intelligence (OSINT) that tends to evolve over time (Glassman and Kang, 2012) hence being able to reflect the changes that have occurred.

4. Analyse and identify devices that SHODAN[5] can use to pull banners from. In addition to this, the research will identify critical variables that can be used to show levels of vulnerability but also those that can be quantified so as to provide a scale which can be used as a guiding principle. Should time permit, it will identify what type of vulnerability or attack each of them can cause of systems and applications.

1.3 Scope and Limits

This research aimed to focus primarily on East African Countries where SHODAN records instances of vulnerability. The study used two forms of open source intelligence data namely SHODAN and Internet background radiation (IBR).

SHODAN is a search engine that crawls for banners on web pages by interrogating ports, web interfaces and pulls off the service banners and indexes them for searching rather than the standard web content (Schearer, 2010). SHODAN is designed to help the user find specific nodes (servers, routers, switches, webcams etc.) with specific content in their banners. SHODAN categorizes the data into specific variables, for example, it gives port numbers, Internet Service Providers (ISP), device type etc which are used for analysis. Thus all these variations will be taken into consideration in order to find the focus of the research.

IBR consists of non-productive data packets on the Internet (Cooke, Bailey, Mao, Watson, Jahanian, and McPherson, 2004) which are addressed to unused IP addresses or ports where there is no network device set up to receive them and often times shows evidence

[5]https://www.shodan.io/

of either malicious activity or misconfiguration be it temporal or permanent misconfigurations (Pang, Yegneswaran, Barford, Paxson, and Peterson, 2004). Given the fact that there are no legitimate hosts in an unused address block (Polakis, Kontaxis, Ioannidis, and Markatos, 2011), then the data collected therein must be a result of misconfigurations, backscatter from spoofed source addresses or scanning from worms and other probings (Cooke *et al.*, 2004; Wustrow *et al.*, 2010). It, therefore, makes IBR an ideal method for analyzing and quantifying Internet security phenomenon

1.4 Document Conventions

Table 1.1 shows style conventions that have been used throughout this research document:

Table 1.1: Document Conventions

Element	**Description**
Mono-space font	Code samples
Italic font	Name of Appendix, Listing attached in the appendix. It also means a specific format of data (e.g .csv) and specific program used for analysis (e.g awk)
Bold font	Header of a table or section. It also means a top variable in a table or a variable that is repeated more than once in a table.
footnote	a note printed at the bottom of a page that gives extra information about something that has been written on that page. It is presented within the text with a superscript number and at the bottom of the page to provide ease of reference. At the bottom of the page is shown as URL
IP address block	Presented as 10.10.10.X with the last octet blinded for security purpose of the owner of the addresses but also accumulation of IP addresses within that block i.e. /24 Netblock
Ports	Ports belonging to TCP and UDP protocols are presented as either 445/TCP or 123/UDP indicating the protocol which they belong to

1.5 Document Structure

The remainder of this document consists of seven chapters structured as follows: Chapter 2 marks the beginning of literature review. This chapter introduces the terms and

concepts used throughout the paper and provides background information on common cybersecurity threats, attacks, and vulnerabilities. Related work is also covered in this chapter and the methodologies previously used. It also provides an overview of the main cyber threat actors and the role that ISPs plays in line with cybersecurity. Lastly, it elaborates how cyber warfare has evolved over the years.

Chapter 3 elaborates cybersecurity from an African perspective. It shows why African perspective is different from that of other continents especially that of the American and European continents. Among reasons highlighted include digital colonization, factors driving the risks, level of security awareness which comes along with ever fast growing access to the Internet. In addition to these, it also eludes to the proliferation of devices and an increase in mobile applications that involve money transactions like mobile banking. Lastly, this chapter addresses the link that exists between cybersecurity and its legal framework. We get to see how far countries in East Africa have gone in addressing cybersecurity threats with their current laws.

Chapter 4 describes the characteristics of the data used in this research. It also justifies why certain characters were picked over others. It explains the limits that come with the choice of using OSINT, efficiency, and legality of using OSINT. It explains the relevance of using OSINT over any other form of data and its ethics. The tools that were used to clean up the data, scripts, and lines of code written and program used are also explained in this chapter.

Information security perspective from Internet Background Radiation (IBR) data analysis is explained in Chapter 5 where related work is done with IBR is given in a brief summary. It goes further to explain and discuss the two components that are key variables to the study in great detail. It also performs a comparative analysis of the variables in all the countries of study which is shown either by tables or graphically presented. The chapter concludes by offering recommendations and justifying why unique IP addresses are significant as an element of study.

In Chapter 6 information security perspective from SHODAN data analysis is explained. This chapter has results in the form of graphs and tables. Given the fact that this study has never been done in the East African Community (EAC), this chapter explains the relevance of these results matching up with one of the objectives set. In Chapter 7 the research addresses the issue of quantifying vulnerabilities by the use of Common Vulnerability Exposures Identity (CVE-IDs) which are put into CVSS framework in order to compute scores. Data presented in Chapter 4 presents two CVE-IDs that are used to

demonstrate this process. This is later followed by the significance of these scores to the EAC.

This research concludes in Chapter 8 with an overview of the results acquired in Chapter 5, 6 and 7 focusing on the key findings that give insight to information security community. In addition to this, it justifies the significance of this study and evaluate the objectives set in Section 1.2 to assess if they have been met and provide possible future work that comes as a follow up to this study.

Chapter 2

Literature Review

This chapter provides background information relevant to the research presented in the remainder of this document. An introduction and definition of cyber security concepts that have been used throughout the research are approached in Section 2.1. These concepts include what constitutes cyber warfare, cybersecurity, cyberspace and what constitutes CII. In Section 2.2 common cybersecurity vulnerabilities are introduced into the study. These include injection vulnerabilities, buffer overflows, broken authentication, session management and security misconfiguration. A brief description of cybersecurity common threats and attacks are highlighted in Section 2.3. It explains what DDoS is, short lists of some of the malicious software that has invaded cyberspace and how botnets are used.

Section 2.4 defines the main cybersecurity threat actors followed by an evolution of cyber warfare in Section 2.5. The role of Internet Service Provider (ISP) in the circles of cybersecurity is explained in Section 2.6, as they are at the center of providing Internet services to their clients. This chapter concludes by looking at work done by other researchers in the same field, the methodologies which they used. It also justifies why CVSS and CVE have been used to quantify the threats, attacks, and vulnerabilities. This is covered by Sections 2.7 to 2.9. The chapter ends with a summary of what has been covered in Section 2.10.

2.1 Definition of Cyber Security Concepts

Discussion in this thesis makes use of jargon common to cybersecurity as such this section lists and describe some of the cybersecurity concepts that have been used in this study.

Furthermore, it explains how some of these concepts are used by cybersecurity actors in order to achieve their objective.

2.1.1 Cyberspace

The notional environment in which communication over computer networks occurs i.e metaphor for describing the non-physical terrain created by computer systems within which people can communicate with one another (Winterfeld and Andress, 2012). It is a virtual environment that has no boundaries and as long as one has Internet connectivity then they are able to use this social space for communication and business transactions (Smith and Kollock, 1999).

2.1.2 Cyber Warfare

These are actions taken by a nation-state, be it attacking or defending, against another country with the aim of penetrating the other country's information and computer networks in cyberspace (Winterfeld and Andress, 2012). A cyberspace is a virtual space meaning these attacks are not done by the use of physical weapons, like vehicles or military personnel, but rather over networked systems (Nicholson, Webber, Dyer, Patel, and Janicke, 2012).

Cyber warfare attacks aim to jeopardize the confidentiality, integrity, and availability of cyber infrastructure (Nicholson *et al.*, 2012). Cyber infrastructure consists of computational systems, data and information management, advanced instruments, visualization environments and people, all linked together by software and advanced networks to improve scholarly productivity and enable knowledge breakthroughs and discoveries not otherwise possible (Winterfeld and Andress, 2012). Attacks can range from website defacements, DoS attacks to the destruction or covert monitoring of data (Singh, Kumar, Sachdeva, and Sidhu, 2012).

2.1.3 Cyber Security

It is a body of technologies, processes, and practices designed to protect insecurities related to networked computers, critical information infrastructure, programs and data from attack, damage or unauthorized access (Hansen and Nissenbaum, 2009). Some of

these technologies include antivirus tools, firewalls, file integrity checkers, vulnerability scanners and intrusion detection systems that protect, detect and react to unauthorized system access (Mukkamala, Sung, and Abraham, 2005)

2.1.4 Critical Information Infrastructure

A term used by governments to describe assets that are essential for the functioning of a society, the well-being of the national and international economy, security and quality of life and is highly dependent on interconnected national software-based control systems for their smooth, reliable and continuous operation (Roman, Alcaraz, and Lopez, 2007). Critical infrastructure encompasses a wide array of physical assets such as the electric power grid, banking, telecommunications, oil and gas pipelines, transportation networks and computer data networks (Gorman, Schintler, Kulkarni, and Stough, 2004).

2.2 Common Cyber Security Vulnerabilities

In this computer era, cybersecurity vulnerabilities are among the top list of concerns for information security managers in big organizations especially those that keep sensitive data and often deal with monetary transactions. Cybersecurity venerabilities refers to any form of susceptibility that exists in computer systems, applications or programs and poses a risk of being exposed to both internal and external attacks facilitated by cyber actors (Banzhof, Cook, Helffrich, and Lawson, 2004). These weaknesses in systems have the potential to distort the confidentiality, integrity and availability of the system's data should they be exploited by the threat actors. In this section, we focus on the most common cybersecurity vulnerabilities that are often found in software, operating systems or services that have access to the Internet or any form of connectivity.

2.2.1 Injection Vulnerabilities

Injection vulnerabilities allow attackers to relay malicious code through an application to another system be it its system calls, shell commands or calls to back-end databases which is done by injecting a written script into poorly designed applications and executing it as long as it has an interpreter (Couture, 2013). Injection occurs when user-supplied crafted input value is sent to an interpreter as part of a command or query thus tricking the

interpreter in the process to executing unintended commands (Couture, 2013; Shar and Tan, 2013). The injection occurs via the input data from the client to the application.

The consequences are particularly damaging as an attacker can bypass deeply nested fire-walled environments and either obtain, corrupt, or destroy database contents as the attacker will have unauthorized and unlimited access to the database (Shar and Tan, 2013). Of these injection vulnerabilities, SQL injection (SQLI) and cross-site scripting (XSS) are the two most common and serious web application vulnerabilities that have been in existence for the past decade (Shar and Tan, 2013). Considering that more than 20 percent of all web vulnerabilities are attributed to SQL injection, making it one of the most dangerous software vulnerabilities, this study has focused more on it than any other form of vulnerability (Couture, 2013; DuPaul, 2016).

SQL injection is a software vulnerability often associated with web application security vulnerability in which an attacker is able to submit a database SQL command that is executed by a web application and exploits security vulnerabilities at the database level (Shar and Tan, 2013). An SQL injection attack can occur when a web application utilizes user input values without proper validation or encoding as part of a command or query (Ali, Shakhatreh, Abdullah, and Alostad, 2011b). Attackers utilize this vulnerability by providing specially crafted input values to the SQL interpreter through a web front-end in such a manner that the interpreter is not able to distinguish between the intended commands and the attacker's specially crafted data (Jang and Choi, 2014). SQL injection allows an attacker to make any form of malicious alteration to the web application including creating, reading, updating, altering or deleting data stored in the back-end database (Jang and Choi, 2014). In its most common form, an SQL injection damages websites and its attacks can cause DoS and gives access to sensitive information such as social security numbers, credit card numbers and/or other financial data (Ali *et al.*, 2011b).

2.2.2 Buffer Overflow Vulnerabilities

A buffer overflow is an anomaly where a program attempts to write more data to a fixed length block of memory (buffer) than it was intended to hold which then forces the extra data to overrun the buffer's boundary and overwrites to the adjacent buffers (Kim, Lee, Han, and Choe, 2010). This extra information that overflows corrupts or overwrites the valid data held in the adjacent buffers creating data integrity issues (Kim *et al.*, 2010). Though it may occur as a result of programming errors, often times it is a result of buffer

overflow attacks in which the extra data may contain codes sending new instructions to the attacked computer to perform any operation of a superuser (Fen, Fuchao, Xiaobing, Xinchun, and Bing, 2012). Buffer overflows ought to be given more attention because they constitute at least a third of all the severe remotely exploitable vulnerabilities (Zitser, Lippmann, and Leek, 2004).

Among the instructions to be carried out include damaging the user's files, changing data, disclosing confidential information or executing arbitrary code that is beyond security policies of the program (Cowan, Wagle, Pu, Beattie, and Walpole, 2000; Alouneh, Kharbutli, and AlQurem, 2013). In critical conditions, buffer overflow puts a system in an infinite loop which leads to a system crash or system availability becoming void and users fail to use the system referred to as DoS (Cowan *et al.*, 2000).

As much as many programs are affected by buffer overflow attacks, the degree to which they are exposed differs greatly with some more exposed than others depending on the compiler and interpreter that were used during development. For example, if a cyber attacker attacks an application designed and written in C or C++ the attacker has a greater chance to fully compromise the entire targeted system than he could with an application developed in Python because C/C++ applications have no built-in protection to buffer overflows (Kim *et al.*, 2010; Fen *et al.*, 2012). Despite the fact that each operating system's platform behaves different from the other, i.e. Windows operating system behaves differently from Linux, every platform in common use today is affect by this vulnerability (Fen *et al.*, 2012; Alouneh *et al.*, 2013).

2.2.3 Broken Authentication and Session Management

This is another critical area where if not properly managed it can bring web application systems to its knees due to its poor management. Often times these vulnerabilities come as a result of negligence to protect one's credentials and session tokens and more importantly failure to adhere to standard security policies that safeguard such systems during their development (Singh and Sharma, 2015). Such habits and behavior include poorly developed authentication and session management protocols for application systems, failure to put limits to log in attempts, letting an application log in without a password or remembering credentials from previous logins (Nagpal and Nagpal, 2014).

Due to their nature, broken authentication and session management attacks have a high risk of compromising the credibility of companies as they cause loss of integrity and

confidentiality of data and act as a gateway to further malicious attacks to the targeted organization (Pannu, 2014; Ahuja, Johari, and Khokhar, 2015). If it is an administrative account that has been affected by a cybersecurity actor, the attacker acquires all privileges that are associated with the account and can cause severe damage to all systems associated with the acquired credentials (Ahuja *et al.*, 2015). Broken authentication and session management vulnerabilities affect all environments be it web servers or application servers (Pannu, 2014) and is ranked second according to Open Source Web Application Security Project Top 10-2013 (William and Wichers, 2013).

2.2.4 Security Misconfiguration

Often times attackers capitalize on things that seem out of order as far as configurations are concerned in order to get access to systems, be it a server, database, firewall or an operating system platform (Eshete, Villafiorita, and Weldemariam, 2011). Any mistake that administrators and software developers make when setting up firewalls, servers and developing systems pose a threat to the system is what is referred to as security misconfiguration (Cuppens, Cuppens-Boulahia, and Garcia-Alfaro, 2005). This could be maintaining default settings like passwords, default ports, using the same password for all super user accounts or failure to limit user privileges in user roles (Nichols and Peterson, 2007). Just like there are tools and algorithms for administrators to detect such misconfigurations cyber attackers also have access to such tools and their sniffing tools make this vulnerability easy to detect (Cuppens *et al.*, 2005; Das *et al.*, 2010).

The impact of security misconfigurations could be fatal: from data loss which may result in expensive system recovery costs to entire systems being compromised that an overhaul of the system becomes the only way out (Eshete *et al.*, 2011). The best approach to safeguarding systems from security misconfiguration is to change default user accounts, patching systems, perform routine security checks, deleting inactive user accounts and disable ports as well as services that are not being used (Wichers, 2013). The list for such misconfigurations cannot be exhausted, but these should never be overlooked.

2.3 Common Cyber Security Threats and Attacks

Once a threat actor finds a vulnerability in a system they aim at exploring it by attacking the system. In this section, the study looked at the common threats and attacks posed

to the information technology infrastructure by different threat actors who use different means and tools in order to gain access to systems and applications. These threats and attacks range from DDoS attacks, malicious software attacks and social engineering to use of botnets.

2.3.1 Distributed Denial of Service

A distributed denial of service (DDoS) attack occurs when the attacker gains access to a machine on a network or system that is not heavily guarded or it is insecure with the aim of gaining total control of the machine, normally referred to as a zombie when it is under the control of the attacker (Vadehra, Chowdhary, and Malhotra, 2015). The attacker will need more than one machine to orchestrate a team of zombies who are forced to flood packets towards the victim's network and transport levels (Vadehra *et al.*, 2015). The overall objective of DDoS is to make the services offered by the attacked system become unavailable to its users or force the system not to serve legitimate users by exploiting software and protocol vulnerabilities (Singh, Kumar, Sachdeva, and Sidhu, 2012). It is relatively hard to identify the source of an attack because, in most attacks, the attacker either spoofs source IP address from the victim to a large set of Internet servers or uses legitimate ports when controlling the zombies making it hard to track down (Hariri, Qu, Dharmagadda, and Ramkishore, 2003). Given the fact that most of these attacks capitalize on insecure ports, vulnerability analysis of the network is key to safeguarding against DDoS (Hariri *et al.*, 2003)

2.3.2 Man in the Middle Attack on HTTP

If communication via computer networking has not been entirely secure, the attacker can easily gain a Man in the Middle (MiTM) position using a technique called Address Resolution Protocol (ARP) spoofing to intercept and modify traffic between victims. With ARP spoofing an attacker sends falsified ARP messages over a local area network with the aim of linking his address (an attacker's MAC address) with the IP address of a legitimate computer or server on the network (Xia and Brustoloni, 2005). Anyone on one's shared media in networks such as Ethernet (with hubs) and Wi-Fi can send a spoofed ARP packets and the victim will unknowingly start sending all its traffic through the attacker instead of the router.

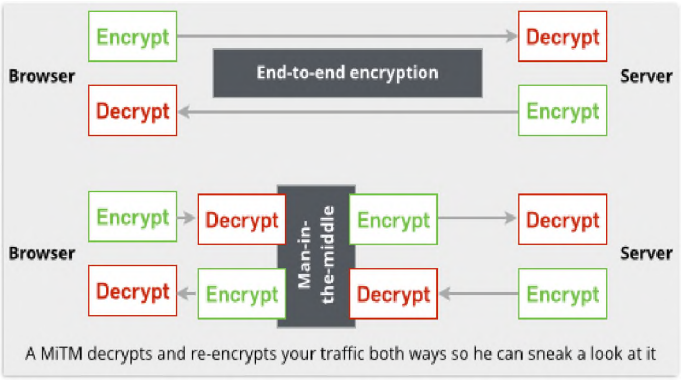


Figure 2.1: MiTM Attack Decrypts and Re-encrypt Schema (Jelic, 2016)

In MiTM attack, a malicious actor eavesdrops on a conversation between two parties by firstly intercepting an initial message sent by the legitimate client, decodes the encryption keys then impersonates the legitimate clients involved and gains access to information that the two parties were trying to send to each other (Nam, Kim, Kim *et al.*, 2010). The attacker can read, modify, inject, or drop any packet, even if client and server authenticate and encrypt all packets even if OpenSSL/SSL is used (Asokan, Niemi, and Nyberg, 2003). Once the decoding of the initial message is done the attacker relays packets each of the two legitimate clients and they assume they are talking to the intended client on the sender or recipient side without their full knowledge of the attack. Figure 2.1 illustrates how a MiTM attacker decrypts and re-encrypts user's traffic both ways (from client to server and vice versa) so as to gain access to the data rendering the client and server authenticate and encrypt all packets as useless.

2.3.3 Malicious Software

Malicious software abbreviated as malware is the underlying platform to cyber-crimes that Internet users are currently faced with: from simple self-propagating code using basic languages like a spam to highly sophisticated ones where developers compete for the most efficient malware to deploy, like DDoS (Lindorfer, Di Federico, Maggi, Comparetti, and Zanero, 2012). This section highlights the commonly used malware subtypes utilized in the execution of cyber crimes in recent years.

Virus

A computer virus is a self-replicating computer program that spreads by attaching itself to executable files and their environment and modifies them without the computer operator's consent (Nachenberg, 1997). A computer virus comes in different forms and types but at the end of the day the damage they cause to systems follows the same pattern: infect an executable file, thereafter trigger pull checks that verify how conducive the environment is to deliver the payload and finally do damage by delivering the payload of the virus (Joshi and Singh, 2013).

Ransomware

A ransomware is a direct revenue generating malicious software that infects a machine with full ransomware payload. Often the malware is usually delivered as an email attachment and any machine that downloads such a file gets infected (Savage, Coogan, and Lau, 2015). Using either crypto or locker payload the attacker denies users access to their devices by locking or prevents users from accessing files or data on their own machine until they pay a specified amount of money in the form of bitcoins demanded by the attacker, then they are emailed a crypto key to unlock encryption used (Everett, 2016).

Spyware

A spyware is a malicious software that is installed on the victim's computer or system to secretly gather information about the user's activities like emails or personal files, login credentials on all applications on the machine or websites by the use of key-loggers or back doors (Abualola, Alhawai, Kadadha, Otrok, and Mourad, 2016). Since Spyware cannot easily spread on their own, they work hand in hand with Trojans who perform the role of masquerading the spyware by offering a legitimate functionality to the user (Piccard and Faircloth, 2006). Although spyware is potentially dangerous in the home-user environment, it becomes a true disaster in the corporate enterprise and government sites and systems as well if such information falls into the wrong hands (Baskin, 2006).

Worm

A worm is an independent and self-replicating program that spreads copies of itself or parts of itself to other computers, commonly across network connections, or remote machines.

These copies are themselves fully functional independent programs, which are capable of spreading further without the need of another program (Fosnock, 2005). Categorised into two groups, direct and indirect worms, they cause damage in the form of information loss, information theft, and DoS attacks making them more dangerous than viruses as they can spread without any human interventions once they infect one machine on the network (Aiello, Avanzini, Chiarella, and Papaleo, 2006).

Trojan

A trojan horse is a self-contained malicious software that often infects a victim's machines by attaching itself to applications or useful downloads that are often executable (Bowles and Hernandez-Castro, 2015). The damage caused by a trojan ranges from loss of intellectual property and finance, deleting user files, changing user settings to capturing typed keyboard characters or enable the creation of a backdoor in the victim's machine to enable a remote user access to the system (Pu, Chen, Cui, Shi, Guo, and Qi, 2013).

2.3.4 Botnets

Derived from the word 'robot' a bot is an executable file or program that controls a network of compromised machines (botnet) by the use of predefined command and control (C&C) protocol (Vadehra *et al.*, 2015). Thus a botnet is an interconnected network of malware-infected machines that are controlled by cyber criminals without the user's knowledge (Stone-Gross, Cova, Cavallaro, Gilbert, Szydlowski, Kemmerer, Kruegel, and Vigna, 2009). The attacker uses a botmaster to remotely control the botnet. Using one of the affected hosts on the infected network system, the attacker explores for more hosts on which to install the bots with the core purpose of increasing his/her botnet (Vadehra *et al.*, 2015).

In recruiting vulnerable end host machines i.e connecting bots to their botmasters, botnets use C&C infrastructure which operates on various network topologies (Royal, 2008; Ollmann, 2009a). Through the use of C&C features, it allows a bot agent to receive new instructions remotely (digital bridge into an organization) as dictated by a remote criminal entity thus making them more lethal and distinct than an ordinary malware (Ollmann, 2009b). This compromised host then can be used as an unwilling participant in Internet crime as soon as it is linked into a botnet via that same C&C making it easy to frame an organization for a crime it has no knowledge of having committed.

Once the botmaster accumulates a lot of bots to form a strong botnet, the attacker is able to commit a number of cyber crimes like illegal shutdowns of other systems, spamming or even orchestrate a DDoS (Vadehra *et al.*, 2015). Due to the deceptive nature of botnets, it becomes relatively hard to detect them even with the most up to date antiviruses (Royal, 2008)

2.4 Social Engineering

Social engineering (SE), in information security circles, is the science of obtaining confidential information or gaining access to computational devices by manipulating legitimate users to unknowingly violate typical policies (Hadnagy, 2010). Information systems are penetrated through the use of social methods. Social Engineering has three dimensions: persuasion, fabrication, and data gathering (Tetri and Vuorinen, 2013). The aim of these techniques which manifest the dimension of persuasion is to get a person to comply with an inappropriate request to make them do something which is against the rules or a set of norms often times of free will, trust or fear (Workman, 2007).

With fabrication dimension, it involves techniques, such as impersonation, name-dropping, jargon, piggybacking or using false ID (Hadnagy, 2010). It also involves providing misleading information or cues on purpose with the aim of creating a false image or scenario of the real situation at hand (Tetri and Vuorinen, 2013). In order to be successful in executing an SE attack one needs to be well informed of his target (Allen, 2006) i.e the systems being targeted, the profile of all employees or targets involved, strengths and weaknesses of the targeted system (making the third dimension) and data gathering which is a prerequisite of the other two dimensions.

2.4.1 Phishing

Phishing is the practice of illegally acquiring information from users in which an attacker attempts to direct users to fraudulent websites by using both social engineering trickery to gain access to sensitive information by impersonating a trustworthy third party (Dhamija, Tygar, and Hearst, 2006). Being a form of social engineering, phishing uses fabrication technique as its dimension in order to perform an attack, often takes the form of an email

that appears to be from a trusted entity, shopping websites such as e-Bay[1] or Pay-Pal[2] or even well-fabricated bank websites (Fette, Sadeh, and Tomasic, 2007).

The email asks for either personal or financial details in order to rectify an alleged problem linked with the targeted user account which users that fall for it are then tricked into downloading and installing hostile software (Wu, Miller, and Garfinkel, 2006). This searches the user's computer or monitors online activities to steal private information. Among the worst outcomes include, but not limited to, fraudulent charges against credit cards, identity theft withdrawals from bank accounts, or other undesirable effects (Jagatic, Johnson, Jakobsson, and Menczer, 2007).

2.4.2 Spamming

The practice of flooding the Internet with many copies of the same message, often times in the form of unsolicited junk commercial electronic mail, that eats up a lot of network bandwidth or worse still distribute malware to compromise more hosts on the recipient's network (John, Moshchuk, Gribble, and Krishnamurthy, 2009). Spammers have acquired more skills and are getting more advanced over time to an extent that their attacks are customised and tailored towards certain targets by the use of bots based on the pre-existing packages that the targets are already using (Difallah, Demartini, and Cudré-Mauroux, 2012). With this sort of mechanism, it is possible for two individuals to receive the same email but end up being affected differently.

2.4.3 Reverse Social Engineering

Reverse social engineering is a unique form of social engineering where the victim unwittingly initiates contact and probes the attacker having been deceived into believing that the attacker is part of a legitimate organization that provides support services to the victim (Ivaturi and Janczewski, 2011). It works on enhancing trust from the victim in order for an attack to occur. Reverse engineering consists of three major steps: sabotaging, advertising, and assisting (Irani, Balduzzi, Balzarotti, Kirda, and Pu, 2011). A good example of this is when an attacker replaces a valid technical support telephone number in an organization with his own thus all calls needing such a service will be targeted to the attacker without the victim's knowledge.

[1] https://www.ebay.com/
[2] https://www.paypal.com/za/home

So when the victim calls for technical support they end up speaking to the attacker thus initiating contact with the attacker as a legitimate service provider without any suspicion. Unlike in a normal social engineering scenario, because the attacker has gained immediate legitimacy and trust from the victim, the hacker can receive much more information in these cases than they would from normal social engineering attacks (Krombholz, Hobel, Huber, and Weippl, 2013). Because of this acquired trust, it enables the attacker to launch a wide range of attacks such as persuading victims to click on malicious links, blackmailing, identity theft, and phishing (Irani *et al.*, 2011).

2.4.4 Short Message Service (SMS) Phishing

Phrased as SMiShing, it is a social engineering attack in which the attacker seeks to direct the text message recipient to visit a website where a user is tricked into downloading a Trojan horse, virus or other malware onto his or her cellular phone or other mobile devices (Ivaturi and Janczewski, 2011). It is similar to phishing except in this case as the name suggests it uses SMS message that is purportedly sent from a reputable source, such as victim's bank asking for personal details as a means of contacting the victim unlike an email in phishing (Chien, 2009). Once the attacker succeeds in getting the necessary bank details they can transfer cash from the victim's bank account to their own.

2.5 OpenSSL Overview

OpenSSL is an open-source cryptographic library that provides a robust, commercial-grade, and full-featured toolkit for the Transport Layer Security (TLS) and Secure Sockets Layer (SSL) protocols used in applications that need to secure communications against eavesdropping or need to ascertain the identity of the party at the other end (Yilek, Rescorla, Shacham, Enright, and Savage, 2009).

2.5.1 OpenSSL Process

The library includes tools for generating Relatively Slow Algorithm (RSA) private keys and certificate signing requests, checksums, managing certificates and performing encryption/decryption (Lenstra, Hughes, Augier, Bos, Kleinjung, and Wachter, 2012). It is often used in Internet web servers, serving a majority of all websites it is licensed under

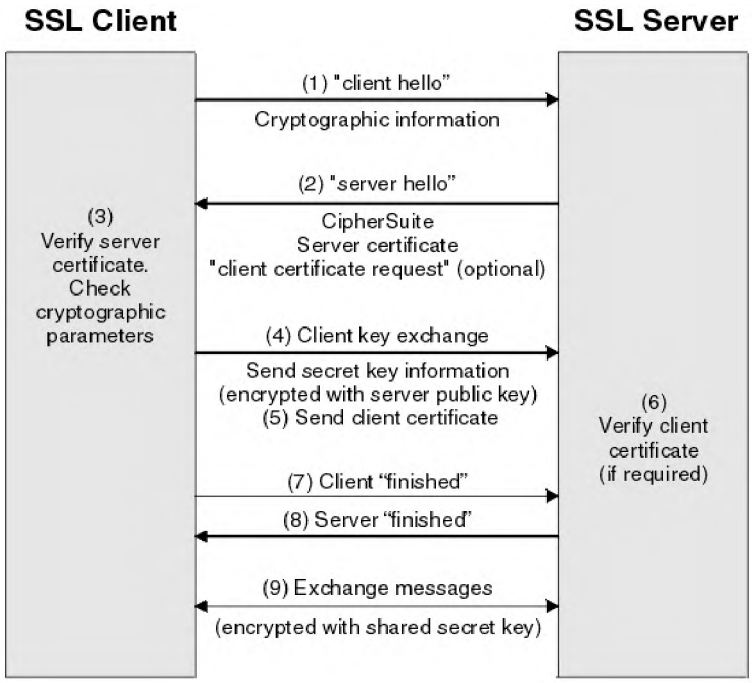


Figure 2.2: SSL Handshake Protocol (IBM, 2017)

an Apache-style license (Arce and Levy, 2003) i.e. you are free to get and use it for commercial and non-commercial purposes subject to some simple license conditions.

Despite the fact that OpenSSL was created to ensure safe communication from one end to the next, its use of handshake phase which authenticates the peers and establishes joint keying material makes it vulnerable to an extent that it allows local and remote attackers to obtain the private key when a client-side communicates with a server (Lenstra, Hughes, Augier, Bos, Kleinjung, and Wachter, 2012). The handshake certificate exchange works on a transport layer (Riccardo and Luca, 2016). During the handshake exchange, the client initiates session to send request *'ClientHello'* to the server and receive response with *'ServerHello'* as shown in Figure 2.2

When server and client exchange certificates to protect the data flowing between them, the communicating peers are located and the certificates are confirmed in their trust key stores, the handshake process is done (Zandbelt, Hulsebosch, Bargh, and Arends, 2008). In certain scenarios, however, the server acts as its own certificate authority and generates an identity certificate that is signed by the server itself (self-signed certificate) whose identity it certifies (Ellison and Schneier, 2000). Such a certificate is just a key carrier: clients cannot verify the server's identity unless they have some independent channel for verifying the certificate.

2.5.2 OpenSSL Vulnerabilities

The bugs present depend on the version of OpenSSL being used among other things. Some of these vulnerabilities include FREAK vulnerability (CVE-2015-0204) (Red-Hat, 2015), Online Certificate Status Protocol (OCSP) stapling vulnerability (CVE-2011-0014) (Chariton, Degkleri, Papadopoulos, Ilia, and Markatos, 2016), Heartbleed vulnerability (CVE-2014-0160) (US-CERT, 2014) and CCS Injection Vulnerability (CVE-2014-0224) (Gutmann, 2014).

OCSP Stapling Vulnerability

The Online Certificate Status Protocol (OCSP) is one of the two ways for obtaining the revocation status of X.509 digital certificate (e.g. SSL & code-signing certificates) and hence maintains the security of a server or other network resource (US-CERT, 2014). OCSP stapling vulnerability occurs when creating a handshake when a client sends an incorrectly formatted (malformed) *'ClientHello'* handshake message, that triggers an out-of-bounds memory access leading to OpenSSL parsing more than the end of the message (Topalovic, Saeta, Huang, Jackson, and Boneh, 2012). In this handshake a client sends an OCSP request to an OCSP responder, which responds whether the certificate is valid or not thus allowing a web server to provide information on the validity of its own certificates rather than having to request the information from the certificate's vendor (Liu, Tome, Zhang, Choffnes, Levin, Maggs, Mislove, Schulman, and Wilson, 2015). This shifts the burden of handling OCSP requests from certificate vendors to web hosts but also more importantly prevents users from transmitting sensitive browsing information to third parties (Chariton *et al.*, 2016). Unfortunately this strength poses a potential privacy risk as it keeps track of user's browsing behaviour to certification authority (Liu *et al.*, 2015).

OCSP stapling vulnerability is present in OpenSSL versions 0.9.8h to 0.9.8q and OpenSSL 1.0.0 to 1.0.0c which in a worst case scenario could allow the attacker to cause a DoS or crash since the parsing could lead to a read on an incorrect memory address (Acıiçmez and Schindler, 2008). In some worst cases mitigating such breaches has been a total failure like the case of Google Chrome who at a later stage decided to disable OCSP permanently (Topalovic *et al.*, 2012).

Heartbleed Vulnerability

Heartbleed vulnerability was disclosed publicly in the first quarter of 2014 by OpenSSL project with versions 1.0.1 through 1.0.1f and OpenSSL 1.0.2-beta containing a flaw in its implementation of the TLS and DTLS heartbeat functionality resulting in severe memory handling bug (Mpofu *et al.*, 2012; Durumeric *et al.*, 2014). With this flaw found in OpenSSL an attacker would be able to retrieve memory intended for applications in chunks of 64KB with repeated exploitation resulting in retrieval of additional memory to further classified information (US-CERT, 2014).

Attackers well equipped with private key knowledge have the potential to perform a MiTM attack against any future communications by intercepting traffic between a vulnerable client and server, altering it according to the attackers preference (Judge, 2014). MiTM occurs where a remote unauthenticated attacker may be able to decrypt and modify network traffic in transit (D'Orazio and Choo, 2016). Other than public HTTPS web services Heartbleed is also capable of affecting servers that make use of TLS, the Tor network, Bitcoin, Android, and wireless networks (Durumeric *et al.*, 2014). Figure 2.3 is an example with full illustration on how an attacker can conduct MiTM as the malicious actor eavesdrops on a client trying to access a bank server and how a malicious actor inserts himself as a relay/proxy into a communication session between customer or system and the bank system. By impersonating the server, the attacker may be able to fool the client into connecting to the attacker rather than the server.

Change Cipher Spec Injection Vulnerability

Like Heartbleed, Change Cipher Spec (CCS) Injection Vulnerability uses MiTM attack to explore the vulnerability found in OpenSSL versions 0.9.8za, 1.0.0m, 1.0.1h (on the client side) and OpenSSL 1.0.1 and 1.0.2-beta1 on servers (Beurdouche, Delignat-Lavaud, Kobeissi, Pironti, and Bhargavan, 2015b). If a CCS message is injected by a MiTM attacker to both client and server right after the ServerHello message both parties weak encryption keys will be computed (Bhargavan, Lavaud, Fournet, Pironti, and Strub, 2014). This weak secret, combined with the public client and server random values, is used to compute the encryption keys on both sides, which are therefore known to the attacker.

The bug is exploited when OpenSSL accepts CCS inappropriately during a handshake (Beurdouche, Bhargavan, Delignat-Lavaud, Fournet, Kohlweiss, Pironti, Strub, and Zinzindohoue, 2015a) i.e there is a small window of opportunity that exists between the time

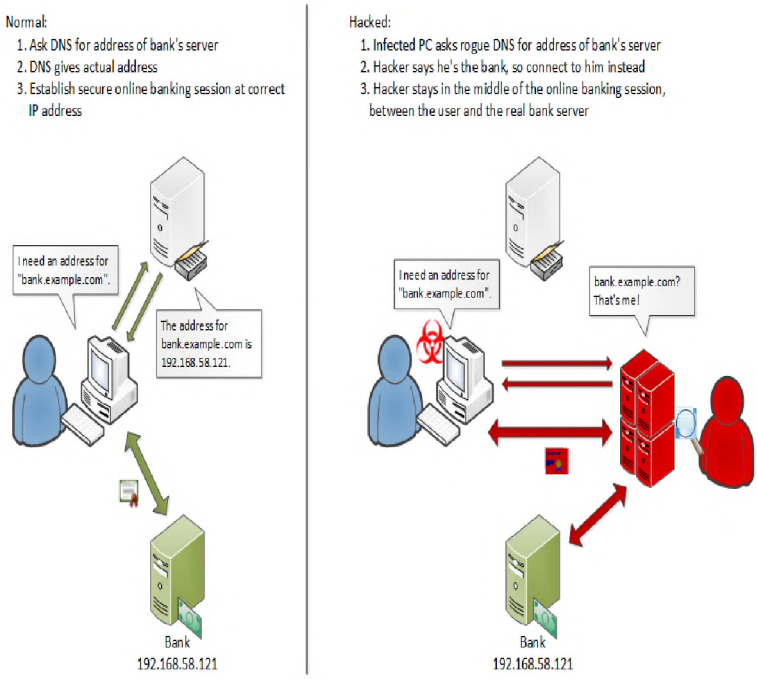


Figure 2.3: Normal Banking Session (left) vs MiTM (right) (Jackon, 2014)

when the first client sends CCS and when the recipients get the CCS and depending on the public key operation being done that gap in time provides window of opportunity for the attacker. Successful exploitation could lead to a security bypass condition where an attacker could gain access to potentially sensitive information, breach of privacy or DoS (Beurdouche *et al.*, 2015b). The upper hand to victims of this vulnerability is that CCS can only be exploited if both server and client are vulnerable to this issue meaning that in the event that one of the two is vulnerable, there is no risk of exploitation (Gutmann, 2014).

These are just some of the vulnerabilities that OpenSSL is prone to. The good news is that patches and updates for all of these have been provided by either the vendors that use such library for encryption or OpenSSL itself. These vulnerabilities are present in all platforms like Windows, iOS devices, Linux systems like Ubuntu, Debian, Red Hat if not properly updated all products associated with these vendors become vulnerable to OpenSSL bugs (Yilek *et al.*, 2009; D'Orazio and Choo, 2016).

2.6 Main Cyber Threat Actors

Classification of cyber threat actors is really a daunting challenge to accomplish but based on the generic motivations of the hacker and the means of cyber attack that these actors commonly use one can come up with categories which may not be applicable to all countries (Sheldon, 2012). In this study, we looked at three categories that accurately captures significant threats, namely: cybercriminals, state-sponsored hackers, and hacktivists. Due to the nature of cyber crimes that occur in the cyberspace it is relatively hard to identify or classify a typical hacker as they can happen from anywhere, but also can be done by anyone with programming expertise or the financial muscle to buy a toolkit that is used to create malware, trojans or botnets (Stillions, 2012).

2.6.1 Cyber Criminals

Cybercriminals are cyber threat actors acting on an individual basis or in groups who commit cybercrimes by the use of computers either as a tool or as a target or as both and are highly motivated by monetary gains from whatever transaction that they carry (Hypponen, 2006). This group of actors is increasing at a faster pace than the other two as the cyberspace now provides a conducive environment for them to carry out their transactions (Hypponen, 2006; Sheldon, 2012). Such activities include stealing of sensitive personal data used to blackmail their owner, financial data, business secrets being stolen from one organization or individual and sold to another. It also includes espionage activities where proprietary information is then sold on the black market especially government secrets (Burden and Palmer, 2003; Nykodym, Taylor, and Vilela, 2005). In other instances cybercriminals who are extremely gifted in developments of cyber toolkit end up selling them to the highest bidder making it easy for even the inexperienced programmer to carry out a sophisticated cyber operation as long as they can afford to buy the tools (Stillions, 2012).

2.6.2 State-Sponsored Hackers

These are actors who pose to have a lot of financial support and also human resource as they have the support and backing of the government to perform a large scale, often times destructive cyber operations on another nation (Sheldon, 2012). Such financial muscle and human resource enables these states to carry out the most sophisticated cyber

attacks and develop a unique malware to aid them in their operations thus making them the most dangerous threat actors in the cyberspace (Brenner and Crescenzi, 2006; Murphy and Murphy, 2013).

Such action may include but not limited to cyber espionage, military operations that demand a particular set of skill in large numbers and cyber operations that are potentially controversial in nature, often triggered by the states need for information on the intentions and activities of other nation-states, terrorists and criminal organizations (Stillions, 2012; Murphy and Murphy, 2013). The fear of being left behind by other states in cyber advancements and being vulnerable to other states attacks forces other nation-states to develop offensive cyber capabilities under the pretext of self-defense (Stillions, 2012).

2.6.3 Hacktivists

These are threat actors who hack into computer networks with the sole purpose of sending a message across to the targeted organization and are highly motivated by either a political, religious or social purpose (Taylor, 2005; Greenberg and Chamberlain, 2012). A hacktivist uses the same tools and techniques as an ordinary hacker but does so in order to disrupt services and bring attention to a political or social cause (Vegh, 2002).

Hacktivists include cyber-terrorists who seek to create acts of terror against the public in order to make a political statement or achieve political objectives by the use of cyberspace (Laqueur, 1996; Vegh, 2002). Cyberspace is preferred primarily because it provides publicity on a global scale and also it makes it easy for recruitment, funding, and training of new recruits (Sheldon, 2012).

2.7 Evolution of Cyber Warfare

Cyber attacks were once categorized as either being syntactic or semantic where the former exploited technical vulnerabilities in software and hardware to commit cybercrime and the latter exploited social vulnerabilities to gain personal information (Choo, Smith, and McCusker, 2007). Recently things have changed where the two forms of attacks are merged and are being referred to as blended attacks i.e. attacks that use technical tools to facilitate social engineering in order to gain privileged information (Choo *et al.*, 2007).

2.7.1 Worm Attacks

One of the first ports that were explored by the hackers in the early 1990's was port 80/TCP (Bayuk, Healey, Rohmeyer, Sachs, Schmidt, and Weiss, 2012) which enables external users to access web services. To prevent massive malware attack the use of proxy servers was introduced whose role is to intercept all user traffic headed for the Internet and then comparing the content of the communication to a set of communication rules established by the organization (Bayuk *et al.*, 2012). If there is a conflict between the traffic and the rules of the company the proxy servers block it.

Cyber crimes have moved from worms, viruses and DDoS to advanced persistent threats (APT), identity theft, phishing, pharming, social engineering, weak crypto exploits, and server key theft motivated by different aspects but have resulted in increased costs, loss of property and lost productivity (Webb, 2013). Everything keeps on evolving every day in the cyber security paradigm due to changes on the devices that we currently deal with and that calls for a change in the approach and effort to be put in combating these threats and attacks.

2.7.2 Toolkits

The inception and emergence of new technologies and their usage coupled with the social media boom, well funded and coordinated crimes have pushed the nature of cyber threats to levels that traditional information security practices are finding it hard to cope with (Maddison, 2013). The emergence of new applications that work on mobile devices and smartphones have also broadened the attack vector. Applications like mobile banking and other payment services (like buying of airtime, water and electricity bills) have resulted in attackers designing malware that aim at compromising credentials on these devices and steal data that is stored in them and are equally vulnerable to threats like botnets (Harris, Goodman, and Traynor, 2012).

One would expect that cybercriminals would just be interested in monetary gains from cyberspace and would thus only target banks and financial institutions, unfortunately, this is not the case. With our continual total dependence on Internet government valuable assets, identified as Critical Information Infrastructure (CII), are also falling victims to cyber attacks. CII that is vital to national security, economic development and public health and safety has also attracted the attention of cyber criminals (Choo, 2010).

Once attackers have access to CII they can manipulate the operations of such systems and in extreme cases disguise the operations to a point that even the administrators for such system will not be able to monitor the attacker's activities (Choo *et al.*, 2007). All they get to see is the effect of a malfunctioning system like high and fluctuating voltage in power plants or high concentration of chemicals in water management systems.

Unlike in the past where it was just isolated skilled individuals, the current threat agents could be anyone from thieves with computer expertise to students, business personnel, and even governments who are supposed to be on the defensive side are also offensive; even to the point of funding such attacks (Rosenquist, 2014). Often times governments get involved either directly or indirectly. For example, if the current cyber activities serve their best interest then they may opt to turn a 'blind eye' to such activities knowing they will benefit from the intelligence gathered (Choo, 2011).

The availability of highly sophisticated toolkits to build malware means that an attacker does not need to have a lot of expertise in programming or hacking skill set to build their own malware (Choo, 2011). Tools like Zeus bot malware creator kit enable anybody who is financially capable to make their own malicious code and torment the cyberspace.

There has been an evolution in email security since the Morris worm attack. Emails have become a new form of attack vector where attackers use them for phishing, spyware, blended attacks and spamming (Burke, 2005). The results are so catastrophic to the extent that the damage only takes minutes to occur and the tricky part is that often times the victims, be it an individual or corporation, do not even know until its too late (Burke, 2005). In dealing with these security issues cybersecurity vendors have created security source code analysis software that is incorporated into source code control systems so as to find bugs and flaws before the software is deployed (Bayuk *et al.*, 2012).

2.7.3 Intrusion Detection Systems

Some vendors have developed systems that observe network traffic destined for web server software and web server response. WAFWs are programmed to detect insecure software as it is being used and blocks attempted exploits in real time (Bayuk *et al.*, 2012). IDS have also been developed to act as the second line of defense to a system's security when normal security measure like authentication and data encryption have been bypassed.

These IDS are designed to alert the system administrators of any action that attempts to compromise any of the branches of the CIA triad of a particular resource (Abraham,

Grosan, and Chen, 2005). They do so by detecting any unusual behaviour picked by the system, normally referred to as a misuse pattern or by keeping a profile pattern of a particular threat that has been dealt with previously (Abraham *et al.*, 2005).

IDS have further been refined to DIDS with the aim of improving their detection accuracy. DIDS are embedded inside intelligent agents and are deployed over a large network and communicate with each other, or with a central server that facilitates advanced network monitoring with one overall objective, to allow early detection of planned and coordinated attacks thus giving the network administrators ample time to take countermeasures (Abraham and Thomas, 2006).

Since IDS only alerts the system administrators when the threat is already in the system, there has been a change from intrusion detection architecture to intrusion-avoidance architecture allowing the system architects to decrease the risk of intrusions (Gorbenko, Kharchenko, Tarasyuk, and Romanovsky, 2012). One key thought to remember is that despite all these countermeasures, none of the threats since their inception have disappeared.

2.7.4 Design Based Threat

Design Based Threat (DBT) is an approach used by security experts that implies that the strength of security protection required by a system should be calculated with respect to a technical specification of how it is likely to be attacked (Bayuk *et al.*, 2012). That way one gets to have the maximum benefit of the security design. The best approach thus to these threats was defined by focusing on the technology vulnerabilities that were exploited rather than on the threat actors (Bayuk *et al.*, 2012). The only shortfall is that the approach focuses on the tools and techniques in use rather than specified as a system requirement.

Just as a business plan is brought for review if an organization is making losses, security strategic planning ought to receive the same scrutiny. Various organizations use different approaches to managing its security goals. One such approach is a defense layer approach. Defense in depth is an architecture where security controls are layered and redundant and a vulnerability in one part of the system will be compensated for by another (Bayuk *et al.*, 2012).

2.8 The Role of Internet Service Providers in Cyber Security

Service quality is an important differentiator in a competitive business environment and a driver of service based businesses. ISPs may benefit from obtaining accurate information regarding their customer's assessment of their brand's delivered service quality; such information may enable service brand managers to formulate appropriate marketing strategies in order to achieve a competitive advantage and long-term profitability (Thaichon, Lobo, Prentice, and Quach, 2014). It is commonly acknowledged that service quality drives customer loyalty and company profitability.

However, when it comes to information security, ISPs need to take an active role in ensuring the safety of their clients when they consider delivery of service (Chindipha and Irwin, 2015). For example, they need to know that it is important to remember that a user of a mobile money application connected to an infected ISP network is far more vulnerable than a user engaging in basic web browsing and using email services as this would tarnish the ISPs relationship with the bank involved (Limbu, Wolf, and Lunsford, 2011). Customers are prone to attribute low risks in purchasing from service providers who are reputable in relation to their security practices which creates a good reputation for themselves.

In addition to this, ISPs are expected to know that they shoulder a lot of responsibility by hosting IP addresses of companies and organization as an attack on them would make a lot of clients prone to the same attack Lichtman and Posner (2006). Their security practices and ethics have to be up to date and make sure their security system must always be maintained, updated and patched from any harm that may be caused on the Internet (Quach, Thaichon, and Jebarajakirthy, 2016). Their employers need to be trained to the highest level in all matters of security as they also provide support to their clients on how to keep away from cyber attacks using assigned security standards Rowe *et al.* (2009).

Most studies suggested security as the most important factor in ethics as it involves financial security. Financial security focuses concern on providing financial information and non-financial security which relates to revealing personal information all of which are critical to the clients of the ISP (Flavián, Guinalíu, and Gurrea, 2006). It is therefore imperative that the ISPs have the security of their clients as its top priority.

2.9 Tools and Methodologies Used to Quantify Vulnerabilities

There are a number of other vulnerability 'scoring' systems managed by both commercial and non-commercial organizations. They each have their merits, but they differ by what they measure. For example, CERT/CC produces a numeric score ranging from 0 to 180 but considers factors such as whether or not the Internet infrastructure is at risk and what sort of preconditions are required to exploit the vulnerability(Mell, Kent, and Romanosky, 2007). The SANS vulnerability analysis scale considers whether the weakness is found in default configurations or client or server systems (Yazar, 2002). Microsoft's proprietary scoring system tries to reflect the difficulty of exploitation and the overall impact of the vulnerability (Mell *et al.*, 2007). While useful, these scoring systems provide a one-size-fits-all approach by assuming that the impact of a vulnerability is constant for every individual and organization.

2.9.1 Common Vulnerabilities and Exposures (CVE)

CVE is a free online dictionary of standardized identifiers made up of a compilation of known system security vulnerabilities and exposures aimed at facilitating the exchange of security-related information across separate network security databases, setting the benchmark and speed up vulnerability analysis of computer systems (Chen, Zhang, and Chen, 2010). The main objective of CVE is to bridge potential gaps in security coverage and a provide common ground when it comes to evaluation of products made from different vendors yet experiencing similar vulnerabilities. CVE entries include CVE identifier number known as CVE-ID which is unique to each vulnerability, description of the vulnerability, references to related products and reports, information about the type of vulnerability, time stamps, and severity scores (Mitre, 2016). A practical approach to this process is shown in Chapter 7.

2.9.2 Common Vulnerability Scoring System (CVSS)

Common vulnerability scoring system is an open framework that is used to assess and quantify the impact of software vulnerabilities as well as the severity of computer system security vulnerabilities (Mell *et al.*, 2007). This is done by scoring the potential exploitation impact of these vulnerabilities and its characteristics. The framework is designed to

rank information system vulnerabilities and provide an end user with a composite score representing the overall severity and risk the vulnerability presents (Houmb *et al.*, 2010). Previously, quantification of software vulnerabilities had its deficiencies, which include failure to rate certain vulnerabilities and putting them on a scale of severity (Mell, Scarfone, and Romanosky, 2006). In addition to this, it was not possible to prioritize which vulnerability to focus on first as far as urgency and response were concerned since some of the vulnerabilities change over time (temporal metrics effect) hence they lose their priority of being threats to systems (Schiffman, Wright, Ahmad, and Eschelbeck, 2004).

Security of data and systems consist of Confidentiality, Integrity, and availability (CIA triad) and each organization based on the industry they are involved intends to prioritize one over the other (Petajasoja, Kortti, Takanen, and Tirila, 2011). For example, banks prioritize availability of their system more compared to military organizations who want their data to be more confined and secretive hence prioritizing confidentiality. It also enables the capability to customize security problems that organizations face and accommodates incompatible scoring systems that produce inconsistent scores (Scarfone and Mell, 2009). These did not come out clearly in the previous methodologies. Its strongest advantage is that it is free, easy to use and understandable thus provides a common way to communicate with various IT managers, cybersecurity experts and vendors (Holm and Afridi, 2015). More details are shown in Chapter 7.

2.10 Summary

This chapter provided context to the nature of this research study and elucidated cybersecurity concepts in Section 2.1. It went further to elaborate more on the common cyber security vulnerabilities and how each of them is explored in Section 2.2. In Section 2.3 the cybersecurity threats and attacks that can occur to vulnerable applications and systems were explained. This was followed by the techniques that are used by the attackers in order to gain access to systems in Section 2.4.

It gave a snippet of some of the terms that were used later like OpenSSL and OpenSSL vulnerabilities in Section 2.5. In Section 2.6 cyber threat actors were defined according to their categories and a brief evolution of cyber warfare was explained in Section 2.7. The role of ISPs in line to cybersecurity was explained in Section 2.8. The chapter closed with an explanation of the tools and methods that can be used to quantify vulnerabilities in Section 2.9.

Chapter 3

An African Perspective to Cyber Security

This chapter provides background to why Africa's perspective to cybersecurity is the way it is. In Section 3.1 it explains the motivation for an African cybersecurity perspective which is attributed to its dependence on automated systems. In Section 3.2, the study goes further in detail to expand the risks that come with this attitude to cyber security i.e. it lists factors that are driving cybersecurity risks and attempt to provide potential solutions for each. In Section, 3.3 the study looks at cyber legislation in an African context by first looking at all the three main regions in Africa. In Section 3.4 the study focuses on cyber legislation in East Africa which defines the scope of this study. The chapter ends with a summary doing a recap of what has been discussed in Section 3.5.

3.1 Motivation for an African Cyber Security Perspective

Just like any other continent, Africa has its own perspective on cybersecurity. Like the rest of the world, the fact that Internet is used by a lot of Africans means that as a continent, it cannot run away from the cons that come with it. Not much has been done to ensure readiness for cyber warfare because very few countries in Africa have digitized their military equipment or CII (Van Niekerk and Maharaj, 2013). Cybersecurity requires the presence of digital technology thus the level of security offered by such technology is proportional to its value. It is not that African countries do not know the significance of

cybersecurity but rather that most countries in Africa do not have CII hence the attention is given to the fast-growing field of Information Communication and Technology (Cole, Chetty, LaRosa, Rietta, Schmitt, Goodman, and Atlanta, 2008a)

Somehow African countries have managed to split the components of security i.e. national security, human security, and economic security and feel that national security is only the duty of the military while the rest focus on other pressing matters such as poverty, corruption, and health-related issues, for example, HIV and AIDS (Harris, Goodman, and Traynor, 2012). Due to the financial constraints and other domestic issues that demand attention from governments, the legal obligation of cybersecurity has been pushed to the private sector (Cole *et al.*, 2008a). More evidence is seen when we realize that less than 10% of the countries in Africa have cybercrime legislation, or even CSIRTs that help to combat unexpected cyber attacks and contain them before they get out of hand (Van Niekerk and Maharaj, 2013). This means that it would take time for them to detect the intrusion and even more so to contain the situation. By that time most valuable assets information would have been accessed.

3.2 Factors Driving Cyber Security Risk in Africa

Africa has not always been the center of cyber crimes but of late there have been a lot of incidents indicating cyber crimes like money scam and identity theft. This section looks at some of the factors that are driving this rise in crimes.

3.2.1 Bandwidth Penetration

Developing with limited to no Internet connectivity has made most African countries and private organization overlook the dangers of cyber threats that are in existence (Harris *et al.*, 2012). As long as Africa continues to turn to ICT for modern-day solutions to solve its problems then it cannot afford to overlook the significance of cybersecurity. To begin with, there has been an increased prevalence and affordability of broadband connectivity in the Sub-Sahara region which is coupled with lack of online security awareness (Chindipha and Irwin, 2015). A drop in broadband costs means more people can afford Internet connectivity and with little awareness, more harm than good could come out of it.

3.2.2 Cyber Security Awareness

In this aging technology is the driving force of a country's economy such that if productivity is to be increased and economies well managed then countries in Africa cannot overlook the significance of safeguarding such systems once they are put in place (Longe *et al.*, 2009). This section looks at how things are currently in line with Internet usage and the threat it poses to the users. Thereafter, it looked at how such threats can be mitigated by using case studies that have been tried before and produced significant results.

The Current State of Affairs

It is a great loss to have a system that is performing well and boosting economies, but cannot protect the information it keeps (McCrohan, Engel, and Harvey, 2010). This alone will ensure competitor's access to privileged information thus brings about the need to raise the essence of approved cybersecurity practices (Thomson *et al.*, 2006). Having the full knowledge that any robust system has one major point of weakness, and that is the human involvement in the interaction with the systems (Thomson *et al.*, 2006), it follows that failure to raise the awareness of people in this regard will render systems more vulnerable to cybersecurity risks.

Cybersecurity is more than just a strong password, strong firewall or keeping an up to date antivirus, yet this is all that a significant amount of people around the globe who use the Internet know (Berti and Rogers, 2004). Proper training and awareness on issues such as data encryption, proper password recycling, the essence of penetration testing through awareness campaigns among other practices ought to be a necessity for each organization especially those that hold sensitive data and perform monetary transactions like banks (Frank, 2008). Cybersecurity threats are not taken seriously especially in Africa hence making cybersecurity awareness a daunting task to accomplish (Cole *et al.*, 2008a). A contributing factor to this perception is due to the fact that most of Africa's CII is not online based such that they can be exposed to attacks making governments in such countries negligent to awareness efforts.

Due to lack of proper cybersecurity awareness from system end users especially those that have e-commerce facilities, they end up being easy prey for hackers who take advantage of their ignorance on the technical know-how of protecting confidential and personal information (Thomson *et al.*, 2006). When the employees observe that the organizational heads and their standards are high they will be motivated to change their behaviour to

meet the current standards (Thomson and von Solms, 2008). Employers and employees in each and every organization need to be informed of their role in the safeguarding of the integrity, availability, and confidentiality of company's assets. They need to be informed of the cause and effect of their actions when negligence accompanies their transactions with the production environment. Well trained personnel in matters of information security in an organization can form an extra layer of protection to systems and thus reduce cybersecurity risks (Thomson *et al.*, 2006).

Social media is one of the key areas that make reconnaissance very easy for cybercriminals as it provides a hub of free information that reveals a lot more than what the owner of the user account thinks they are revealing. The social media sites include blogs, collaboration sites (e.g Wikipedia), photography, presentation sharing sites, video-sharing sites and social networks like Facebook[1] and Linked-In[2] all give out a lot of information and in the right hands, one can come up with peoples profiles and cause harm to them or their organizations. This goes back to the point addressed in Section 3.1.

How to Mitigate Cyber Security Risks with Cyber Awareness Campaign

In the same way that muscle memory works in order to return information automatically is how security awareness works, repetitive exposure of information to computer users regarding security risks brings about a significant change in people's conduct (McCrohan *et al.*, 2010). The study conducted by Kevin F. McCrohan, Kathryn Engel and James W. Harvey in 2010 showed two groups whose only difference was the amount of information exposed to the participants of the study. The more computer users are exposed to the dangers of cybersecurity threats at hand and harm they can cause in disrupting normal running of an organization, the better the chance such an organization has at lowering the threat levels.

The study went on further to prove the significance of a strong password. As much as a strong password does not completely guarantee user accounts or application safety, it does offer a high level of resistance against cyber threats. Once people are made aware of the impact that comes when an attacker targets people's password in order to access sensitive data, it will take a long time before they ignore the value of the robust password. In the long run, it will minimise the risks posed to their information.

[1] https://www.facebook.com
[2] https://www.linkedin.com

Those who have experienced great loss due to such threats tend to jump at such opportunities and are ready to implement the new practices of safeguarding information as it is evidenced in the study conducted in 27 different countries for both governmental and nongovernmental organizations (Symantec, 2010). A study conducted in 2005 showed that conducting information security awareness campaigns has a significant impact in enabling an organization to minimise security risks that come about with the use of Internet applications and systems (Bulgurcu, Cavusoglu, and Benbasat, 2010). This means that all organizations that take an initiative in conducting such campaigns have an upper hand in defending themselves compared to those who have never taken any initiative.

3.2.3 Proliferation of Devices

In recent years, Africa has experienced significant growth of wireless technology and use of Internet especially in mobile phone usage (Cole *et al.*, 2008a). Being the leading Continent in cellphone subscription usage (Cole *et al.*, 2008a), and with a lot of development in mobile applications, we can be assured that once more malware has been developed targeting mobile applications especially those with monetary transactions then this will cause Africa to be affected negatively. The worst part is that most African countries have inadequate public sector security professionals, laws and institutions to confront such cybersecurity threats (Harris, Goodman, and Traynor, 2012).

As a continent especially in the Sub-Saharan region, it has grown overly dependent on the Internet in comparison to two decades ago (Longe *et al.*, 2009). The growth in Electronic transactions used to do business, also referred to as E-Commerce, is another reason why active participation in coming up with countermeasures to combat cybersecurity threats is a necessity for Africa (Winn, 2013). Earlier phones that came in Africa did not contain a lot of accessories and complicated functions that require access to the Internet but now most of the phones are smartphones with a lot of complex functionality. Such advancement ought to be coupled with strong security measure because they have become conducive to security threats.

Mobile money applications have made mobile devices like phones and tablets more attractive to attackers. With many of these in abundance in Africa it is no surprise that mobile banking received a warm reception since its inception and have become Africans' second nature because of the positive impact it has in their lives, like ease of doing business or money transfer in places where banks aren't present (Lawack, 2012). Such developments though good, invite a lot of security threats if the user is not well equipped with safety

measures which are a common case in Africa at the moment. From 2003 to 2016 mobile phone companies have seen their customers increase in numbers by at least five times, its growth rate is faster than any other continent at present and at least 30% of its population has mobile phone access within a household (Aker and Mbiti, 2010).

Such big changes in the mobile industry accompanied by the change of ICT infrastructure with equipment like increased network connectivity (both cable and wireless), computers and the coming in of Internet of things has changed Africa's threat landscape. Not only does it have to worry about with issues of inadequate electricity and energy but also how they can safeguard themselves against cyber attacks. What used to be a backward continent in as far as digital race goes is slowly catching up with the rest of the continents since many of the African states have realized the value that improved ICT infrastructures has on their economy but also even in the health sector. Such being the case there is urgent need to catch up in efficient use of it as it might create a hub of cybercrime in the near future with a massive influx of digital equipment that needs Internet access.

3.3 Cyber Security Legislation in Africa

All electronic data transactions need to be subjected to the law and policies for fear of being abused by either party thus promoting technological neutrality in its usage as well as building confidence to all parties involved (Gereda, 2006; Longe *et al.*, 2009). Only a few countries have legislation governing these sort of transactions in Africa.

Of all the parts of Africa, Northern Africa's ICT is the most rapidly developing in comparison to the rest of Africa (Cole *et al.*, 2008a). With this development, they have started implementing some of the cybersecurity initiatives like National Public Key Infrastructure (PKI), Computer Emergency Response Team (CERT/CSIRTs) and Cybercrime Legislation with countries like Tunisia having a feature like CSIRT, PKI and cybercrime legislation (Cole *et al.*, 2008b).

EAC which comprises of Kenya, Uganda, Tanzania, Burundi and Rwanda has been the driving force behind cybersecurity legislation to its members in East Africa (Cole, Chetty, LaRosa, Rietta, Schmitt, Goodman, and Atlanta, 2008a; Venter, Wangwe, and Eloff, 2009). Furthermore, the EAC perceive information security issues as one of the key roles of government agencies and despite reaching its full potential, it has made some progress towards ensuring high levels of cybersecurity; though it needs support from its citizens (Venter *et al.*, 2009).

Other countries outside the EAC like Mauritius and Zambia also have cybercrime legislation (Cole *et al.*, 2008a). Perhaps the formation of EAC is the best approach to the idea of coordination motivates other members to do the same. This is evidenced by the fact that 50% percent the members of EAC have cybercrime legislation while the remaining ones within it are making progress towards the same goal. It is good news that some of the countries that this research will be dealing with have taken initiatives towards combating cybercrime.

In West Africa, Liberia, Nigeria, and Niger have cybercrime legislation (Cole *et al.*, 2008a). Nigeria has taken a step further by forming a cybercrime unit due to the fee fraud scams it was experiencing (Longe *et al.*, 2009). Ghana also has a law enforcement branch dedicated to cybercrime and has passed a comprehensive data protection bill establishing user's rights of data access, control, and consent of use (Harris *et al.*, 2012). Central Africa has the lowest Internet availability on the continent and thus hinders cybersecurity initiatives (Cole *et al.*, 2008a). In Southern Africa, Namibia and South Africa have cybercrime legislation. South Africa has advanced ICT infrastructure giving more explanation as to why they ought to be active in the region (Gereda, 2006).

3.4 Cyber Legislation in East Africa

One of the best ways to combat cybercrime in Africa is to introduce cyber laws that would help to keep in line criminals who terrorize others in the cyber world. Cyber law's core purpose is to ensure proper governance of any form of technology that uses the Internet with the objective ensuring fairness among all parties involved (Vere, 2009). In this section, the paper looks at the cyber laws that have been passed or are being processed in the four countries under study; namely Malawi, Kenya, Uganda, and Tanzania.

3.4.1 Malawi

The Malawian government announced in March 2016 the drafting of new cyber laws to control the use of social media across the country aimed at prosecuting cybercriminals. This decision was reached upon meetings between government representatives and telecommunication companies who reached a consensus of coming up with a national cybersecurity strategy given the rise of cyber crimes lately which leaves the country vulnerable. As a country, Malawi recognizes the role of ICT in economic developments as it

is stipulated in its growth and development strategies (Bande, 2011). They are also aware that these benefits are at times overshadowed by the dominance of cyber crimes in recent years.

Among such crimes include child pornography, cyberstalking, hacks on government sites, human traffic instigated in cyberspace and digital frauds. It also includes consumer protection and deliberate interruption of communication caused by advanced users of the cyberspace which aim at disrupting government operations (Bande, 2013). The government is working on passing the Electronic Transaction Bill (e-Legislation Bill) into the national assembly Bande (2013). This e-Legislation Bill accommodates issue like the establishment of CERT and other laws that govern business transactions making the country a safer environment for investment (Bande, 2013). The positive side of this is the willingness of the government to address this as is indicated in their 2003 Information and Communications Technologies Policy.

3.4.2 Kenya

Kenya, on the other hand, has been a step ahead of Malawi in as far as the passing of the electronic bill is concerned. Among the four countries under study Kenya has developed a lot, technology-wise, which puts them at a high risk should they be attacked. According to Kenya's Information and Communications Act of 2009, they established a body known as the National CERTs as the immediate task force to combat cybercrime (Makulilo, 2013). This task force has full support and funding of the government which makes them good custodians of Kenya's cyber laws and offer training to other bodies when the need arises.

By 2014 they were working on the cybercrime bill while doing a compatibility test with the International stands that promote fairness to all people in the cyber world. Currently, they have the cybersecurity and protection bill as recent as 2016 (Gazette of the United Republic of Kenya, 2016). The cybersecurity and protection bill regulates online content, defines the roles of Internet intermediaries, surveillance of communication, unauthorized access to computer data, access with intent to commit crimes, damaging or illegal disrupting services, cyberstalking, child pornography but also prejudices that are to be followed in order to bring offenders to justice (Gazette of the United Republic of Kenya, 2016).

Realizing the border-less nature of cyber crimes, Kenya's government works with International Criminal Police Organisation (INTERPOL) to combat cybercrimes (Magutu,

Ondimu, and Ipu, 2011). This is to say that should a cyber criminal seek refuge in Kenya and is being sought after by INTERPOL, the Kenyan government gives the INTERPOL[3] all the support they need while at the same time they gain knowledge and skill in gathering digital forensic evidence, detection, and investigation techniques.

3.4.3 Uganda

As of 2005, Ugandan had zero records of either cybercrime or complaint again cyber threat actors. This did not mean that Uganda is immune to cybercrimes as a study conducted later proved this notion to be wrong with evidence coming from different areas (Tushabe and Baryamureeba, 2005). It even had three proposed cyber laws in 2004 all of which spend a lot of time just being reviewed i.e. Electronic Signature Bill ESB), Electronic Transaction Bill (ETB) and Computer Misuse Bill (CMB) (Blythe, 2010).

A lot has changed since then as is evidenced by the digitization of Uganda and the proliferation of devices over the past decade. Unfortunately, like many other countries in Africa, Uganda's cybercrime bill is not given the priority it deserves so that it can be approved to become an Act. One of the major challenges in Uganda is that despite the rise of cyber crimes only less than 12% of them are reported (Baryamureeba and Tushabe, 2004). The number goes further down to less than 3% of them are processed and punished for their crime.

3.4.4 Tanzania

Tanzania took an approach that has not been taken by any of the four countries under study. Instead of addressing first the industrial sector, it introduced courses and program that bridge usage of computers and law in one of its universities (Mambi, 2010). A review of the old laws was done to see if they address changes happening in the cyber world which led to the approval of digital evidence being used in courts. As far as prosecuting cybercrimes this was a significant step although a lot is still expected to explicitly address crimes that are done online. Despite the fact that this step was taken, as of 2012, Tanzania had no cyber laws which meant that any form of cybercrime committed the attacker had a better chance of evading punishment (Saini, Rao, and Panda, 2012).

[3]https://www.interpol.int/Member-countries/World

Another study conducted in Tanzania proposed the training of judges, lawyers and law enforces as a means to enforcers safety in the cyber world since a majority of these prosecutors and law enforcers may not have had this issue in mind during their education (Liganga, 2012). Like in many African countries, Tanzania is failing to cope with the fast-paced changes that are happening in the cyber world with its first laws passed as early as 1980 being enforced with disregard to Internet-related crimes (Mambi, 2010). Although Evidence Act was passed it did not directly deal with cybercrimes thus making the current laws that Tanzania has inadequate to handle cyber-related crimes. In 2015, a cybercrime act was passed in Tanzania which addressed cyber-related crimes such as data espionage, computer-related fraud, illegal data interference, child pornography, cyberbullying among others (Gazette of the United Republic of Tanzania, 2015). Though this is the case a majority of the country seems to oppose it as it infringes on many people's rights and is perceived as draconian with others opting to challenge it in court (Goitom, 2015).

3.5 Summary

This chapter focused on explaining the perspective that Africa as a continent has in regards to cyber security. It explained the motivation behind the notion it holds in Section 3.1 and in Section 3.2 it gave out factors that are a driving force to the rise of cybersecurity threats which are a result of the perception it currently holds. It also explained that most of its CII are not accessible making the governments negligent than how they could have conducted themselves had it been that their infrastructure was exposed to cyber threats. The factors driving the cybersecurity threats are accompanied by proposed means to mitigate. The research also looked at how Africa is standing with regards to cybersecurity laws that ensure proper governance of cyber lawbreakers. This is explained in Section 3.3. This first looked at all African countries in general from the northern region to south. Then in Section 3.4, it narrowed down by focusing on EAC which consists of Kenya, Malawi, Tanzania, and Uganda.

Chapter 4

Data, Data Cleaning and Data Characteristics

This chapter begins with an introduction to Open Source Intelligence (OSINT) data, its background and related work in Section 4.1 since it is the form of data that has been used throughout the study. This is followed by the legal aspect associated with OSINT in the subsection of Section 4.1 and the shortcomings of OSINT. In Section 4.2 the research explains the first form of OSINT which was used in the study i.e Internet Background Radiation (IBR) data. Section 4.2 shows the timeframe which each of the three datasets used in this research was collected. This is followed by how the data was cleaned up and the tools used for processing it in Section 4.3, then in Section 4.4 data characteristics used in the study are explained.

This is followed by Section 4.5 which explains the second form of OSINT used in this study i.e. SHODAN which is linking with a brief introduction given in Chapter 2. In Section 4.5, sample data given in JSON format is shown and a sample of the records collected from SHODAN website is also given.This is followed by Section 4.6 which explains how SHODAN was cleaned up and Section 4.7 elaborates more on the identified variables and characteristics used in Chapters 6 and 7. Scripts used in the data processing are contained in Appendix A. The Chapter closes with a summary in Section 4.8.

4.1 Open Source Intelligence

Open Source Intelligence (OSINT) has been used extensively in a variety of fields since its inception. Among such fields include the business domain where it is used for market analysis with the aim of offering competitive products by identifying weaknesses in competing products (Calof, Wright, and Fleisher, 2008). It is also used by different governments and states in reaching certain critical decisions like stopping a terrorist attack or taking a stand on military action for example where one country engages another based on what they have gathered from unclassified data (Gibson, 2004). It has also proven an effective means of avoiding, containing or even recovering from a disaster be it system related, disease outbreak or disasters caused by weather changes (Backfried, Schmidt, Pfeiffer, Quirchmayr, Glanzer, and Rainer, 2012). In recent years the cybersecurity field has also benefited from the growth of potential information sources in open source data sets like user preference to system interfaces for example (Gibson, 2004).

As it continues to evolve, all these and many other isolated cases show how significant OSINT has been throughout the years (Calof *et al.*, 2008). Defined as any unclassified intelligence that is collected and acquire through the use of publicly available resources with the purpose of addressing a specific scenario or problem, OSINT has become part and parcel of big intelligence gathering organizations to an extent that they cannot imagine life without it (Steele, 2007). Later on in this chapter, the study will look at cases that change over time where OSINT is the only ideal intelligence to consider.

OSINT was preferred in this study because of several factors. Among them include its wide-ranging capability to allowing different inputs, hence providing room for evolution as technology standards evolve (Ponder-Sutton, 2016). Furthermore, it is openly and legally accessible to the public and shareable making all legal implications in its acquisition of no concern (Schaurer and Störger, 2011). It is also relatively cheaper than collecting information via classified means (Pallaris, 2008). More importantly, it is reliable as most agencies of the government use it in making decisions and policies but also provides awareness in understanding global security agenda (Pallaris, 2008).

4.1.1 Legality of OSINT

Focusing on OSINT both the legal aspect of it and it's limitations. Despite the fact that OSINT is acquired from publicly available resources it does not mean that its data can be

processed without ethical standards and considerations (Cuijpers, 2013). A good example is personal data privacy rights and data protection regulation. Personal data is treated as sensitive data if it includes but not limited to health or sex life of a person involved, ethical origin, religious and philosophical beliefs (Best and Cumming, 2007).

In such instances, the country from which this data is collected needs to exercise a high degree of caution before giving it up. Issues like the objective of using the data, how long it will be kept in storage and if it may be used later on for a different purpose (Cuijpers, 2013). All these details need to be clearly defined when collecting OSINT to ensure that it is following all legal requirements before it is used for only consent is not enough.

4.1.2 OSINT Shortcomings

OSINT could also at times present an imperfect picture which contradicts with the actual picture on the ground where real-time data picked by the media fails to fully represent the entire picture at hand, especially if political influence is involved (Pringle, 2003). This makes it relatively hard for OSINT to offer 100% solution as some of the information needed may be missing; thus it calls for supplementary sources to back it up. It can also be misleading at times due to other interference as was the case with Port Harbour incident in 1941 (Pringle, 2003). Though there are such inconsistencies in presenting an imperfect picture, sometimes OSINT provides the best picture of the situation at hand (Swart, 2015).

At times the data that is required to perform specific operations or the assigned task at hand may not be readily available for use as it was the case with this study. At times it is available but the tools required to extract the data may be unavailable. If they are available the user in need of the data has to have a set of skills in order to operate the tools in question or else hire someone who has the skill set needed to filter the data and cleaning it up so that one can make sense out of it as it was also the case in this study especially when working with IBR data sets. When dealing with sensitive information or a study that is of critical nature, OSINT needs verification from a classified source which if not available may delay the results, like data concerning the outbreak of an unknown epidemic for example (Bean, 2007).

Table 4.1: Time-frame for Data Collection

Netblock	Start Data	Stop Date	Total packets
146/8	20-08-2009	17-02-2015	78,562,470
155/8	02-07-2013	17-02-2015	59,314,201
196/8	30-01-2009	17-02-2015	191,206,784
Total			329,083,455

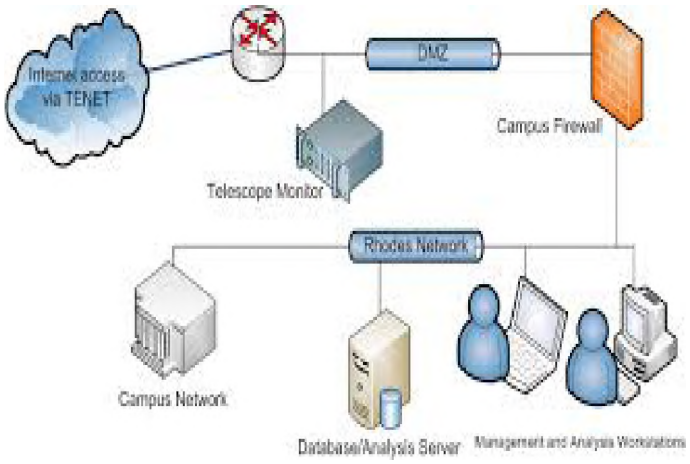


Figure 4.1: Rhodes University Network Telescope Setup (Irwin, 2011)

4.2 Internet Background Radiation (IBR)

The data used in this research was obtained from three of Rhodes University's network telescopes (Irwin, 2011) which had data spanning for a period of over 72 months: from 30th of January 2009 to 2nd of February 2015. A summary of the time stamp and the total packets received by the network telescope is shown in Table 4.1

All the network telescope sensors that were used are physically located in South Africa. Each of the three telescope's sensors consisted of a /24 netblock comprising of 256 individual addresses each, routed to a collection server. The traffic was processed to isolate packets with source addresses from Kenya, Malawi, Tanzania and Uganda. A sample network telescope setup that was used is shown in Figure 4.1.

Furthermore, the blocks of IPv4 space being monitored were distributed across three distinct top-level IP version 4 network address blocks – 146/8, 155/8 and 196/8. As seen in Table 4.1, Of the three network address blocks, 196/8 contained the largest share of the packets followed by 146/8.

Although IBR originates from many sources which can be traced back, there is lack of control over who sends it, when they send it and how many packets they send. This is the

case because of the unidirectional nature of the unsolicited traffic i.e from the attacker to the recipient.

4.3 Data Cleaning

Data cleaning and preprocessing were performed using a combination of common UNIX text processing tools such as *awk*[1], *tcpdump*[2] and *tshark*[3] scripts. In addition, a series of custom python scripts were used to perform the analysis and plotting of data. All scripts used in this preprocessing and data cleaning have been appended to *Appendix A*. The data from network telescope sensor systems was collected using libpcap format, resulting in *.pcap* files being available for analysis.

A series of python scripts were used to pre-process the pcap data and convert it to a *.csv* file format which made it easier to work with when plotting charts and graphs. *Listing 3* in *Appendix A* has more details on how this was brought about. *Tshark* and *tcpdump* were used to filter through the packets from across the globe to IP address range for countries under study with the purpose of getting the network traffic from EAC. Data was further processed using more python scripts along with *awk* to select specific fields of interest. Listing 4.1 show a sample of SHODAN's unprocessed data

4.3.1 Tcpdump and Tshark

Data that was collected by the Rhodes network telescope contained global IP addresses and as such, it needed to be filtered so as to work with only that which was of interest to this study. A *tcpdump* script was written using a specific IP address range of the given countries together. A sample of this filtering script named *Listing 1* is shown in Appendix A. The output file of Listing 1 was used for further analysis in Chapter 6. This script was run in all four countries to analyze data that was collected over a period of 72 months.

Following this script was another script in *Listing 2* (in Appendix A) used to execute the packet filtering script. It was a case of one script being executed by another script where the new file produced was a filtered version of the original script (containing only Kenyan data). This script too was executed for all the four countries but for demonstration

[1]https://www.tutorialspoint.com/awk/
[2]http://www.tcpdump.org/tcpdump_man.html
[3]https://www.wireshark.org/docs/man-pages/tshark.html

Listing 4.1: Sample of Shodan data in JSON format

```
{"_shodan": {"options": {},"id":"8a054271-0a9f-4f0f-a7a1-1f64361
167ad", "module": "http", "crawler": "122dd688b363c3b45b0e7582
622da1e725444808"}, "product": "Microsoft IIS httpd",
"isp": "JHPIEGO", "hash": -1197266993, "asn": "AS37247",
"title": "IIS Windows Server", "ip": 3288433537,
"org": "JHPIEGO", "deprecated": {"opts.robots":
{"new": "http.robots", "eol": "2016-07-01"}, "opts.pem":
{"new": "ssl.chain", "eol": "2016-07-01"}, "html":
{"new": "http.html", "eol": "2016-07-01"}, "opts.sitemap":
{"new": "http.sitemap", "eol": "2016-07-01"}, "title": {"new":
"http.title", "eol": "2016-07-01"}}, "os": null, "cpe":
["cpe:/a:microsoft:iis:8.5", "cpe:/o:microsoft:windows"],
"data": "HTTP/1.1 200 OK\r\nContent-Type:
text/html\r\nLast-Modified: Thu, 12 Nov 2015
19:52:46 GMT\r\nAccept-Ranges: bytes\r\nETag:
\"63ecbcb3831dd11:0\"\r\nServer:
Microsoft-IIS/8.5\r\nX-Powered-By: ASP.NET\r\nDate: Fri,
05 Aug 2016 11:58:26 GMT\r\nContent-Length: 701\r\n\r\n",
"version": "8.5", "location": {"city": null, "region_code":
null, "area_code": null, "longitude": 38.0, "country_code3":
"KEN", "country_name": "Kenya", "postal_code": null,
"dma_code": null, "country_code": "KE", "latitude": 1.0},
"timestamp": "2016-08-05T11:58:29.540646", "domains": [],
"http": {"redirects": [], "title": "IIS Windows Server",
"robots": null, "server": "Microsoft-IIS/8.5",
"host": "196.1.131.129",src=\"iis-85.png\" alt=\"IIS\"
width=\"960\" height=\"600\" /></a>\r\n</div>\r\n</body>
\r\n</html>", "location": "/", "components": {}, "sitemap":
null, "html_hash": 1138219898}, "hostnames": [],
"port": 80, "transport": "tcp", "ip_str": "196.1.131.129"}
```

Table 4.2: IBR Dataset Overview

Netblock	Kenya	Tanzania	Uganda	Malawi	Total Packets
146/8	25,868	11,255	1,068	535	38,726
155/8	27,766	7,473	506	95	35,840
196/8	98,758	37,118	43,543	4,171	183,590
Total	152,392	55,846	45,117	4801	258,156

purposes, this study has shown the Kenyan one. Note that the network telescope collected the data in pcap (packet capture) format as it consists of an Application Programming Interface (API) for capturing network traffic. To work with it there was a need to convert it to CSV format which was done by the script called *Listing 3 (in Appendix) A* as it was a tshark script running for Kenya data.

After filtering from the three net-blocks, the data sets used in this study comprised of 258,156 packets, with Kenya contributing about 59.03 % of the total packets collected. A summary of the data sets collected from EAC is shown in Table 4.2. The packet count for EAC represented only 0.078% of the total traffic received by the three network telescope.

4.3.2 Awk

Already embedded in most of Linux systems, awk is a data-driven scripting language and text processor that can be used for data extraction, reporting and to manipulate text data (Robbins, 2015). *Awk* operates on a line-by-line basis and iterates through the entire file and this makes it more useful for handling text files that are formatted in a predictable way like *.json or .csv* files for example. *Awk* in this study was used to identify specific fields of interest by reading the content of the files it was subjected to a line by line so as to know which fields are necessary for analysis.

4.4 IBR Data Characteristics

The procedure was done to come up with traffic compositions based on the dominant source TCP ports that were captured. This was done in order to establish the intentions of the connection requests (Pang, Yegneswaran, Barford, Paxson, and Peterson, 2004). Since the data spanned over a period of 72 months, a time analysis was done to gauge the number of packets over time and see where the pattern is pointing towards. Given the fact that most of the traffic consisted of TCP packets, the data was further filtered to contain

Table 4.3: Packets by Protocol

	Protocol					
Country	**TCP**	**%**	**UDP**	**%**	**ICMP**	**%**
Kenya	152,392	59.03	23,286	48.95	1,574	97.52
Tanzania	55,846	21.63	23,316	49.01	40	2.48
Uganda	45,117	17.48	969	2.04	0	0
Malawi	4,801	1.86	0	0	0	0
Total	258,156	100	47,571	100	1,614	100

Table 4.4: Number of Unique IP Addresses in EAC

Country	**# of Packets**
Kenya	8,587
Tanzania	3,124
Uganda	900
Malawi	178
Total	12,789

no ICMP or UDP packets in it because their composition was minimal when compared to the total traffic especially in Uganda where there were about 1000 UDP packets and no ICMP packet recorded in Malawi. Table 4.3 shows a summary of the packets by the protocol for each of the countries. Thus for uniformity UDP and ICMP were filtered out.

Another variable that was studied in IBR data set was the source IP addresses. This variable indicated where traffic was coming from in the EAC. Table 4.4 shows a summary of all unique source IP addresses observed in the traffic. Unique IP addresses are distinct and are not repeated anywhere else. In Chapter 5 the study will show how each of these unique IP addresses contributed to the packet count in the study. As in Tables 4.2 and 4.3, Kenya had the most numbers of unique IP addresses while Malawi had the least number of them.

4.5 Shodan

Sentient Hyper Optimized Data Access Network (SHODAN) was used to evaluate the degree to which potential vulnerabilities exist in the publicly accessible infrastructure of the countries in question. SHODAN's primary objective is to collect data from the available ports unlike crawling a Website to access content (Schaurer and Störger, 2011). With Internet access, SHODAN can access all devices that are connected to the Internet

Listing 4.2: SHODAN queries

```
1 Query [1]
2 country:ke,tz,ug,mw "default password"
3 https://www.shodan.io/search?query=country%3Ake%2Ctz%2Cug%2Cmw+
4         %22default+password%22+router
5
6 Query [2]
7 country:mw,ke,tz,ug vuln:CVE-2015-0204,CVE-2014-0160
8 https://www.shodan.io/search?query=country%3Amw%2Cke%2Ctz%2Cug+
9       vuln%3ACVE-2015-0204%2CCVE-2014-0160
```

like Internet-connected cameras, Printers, traffic lights, medical devices, computers, power plants among others (Hill, 2013).

4.5.1 Shodan Data

To come with the data for this study certain queries were run to produce intended variables. Listing 4.2 shows a sample of the queries that were run. While Figure 4.2 shows a screenshot of the output after running the vulnerability query. Figure 4.2 shows a machine in Kenya infected with Heartbleed.

Query 1 for default password was run in November 2016 and it produced 401 records all of which were routers. Out of the 401 records produced, 255 belonged to Kenya representing 63% of the data. Table 4.5 shows a summary of the outcome of the query. Query 2 for vulnerabilities produced 491 records out of which 381 was FREAK vulnerability given an ID of CVE-2015-0204[4]. While CVE-2014-0160[5] which is an ID for Heartbleed produced 110 records. Table 4.6 shows a summary of these records and how each country contributed to the total. Due to the fact that more than one dataset was used in this study each of the datasets was named to make it easier to refer to them. Table 4.7 shows a summary of all data sets used in this study from SHODAN. Table 4.7 also shows the name of the dataset, period which the data was collected shown by Month/Year, description of the data and the number of records that each data set had. Datasets A and B were collected in June 2015 while data sets B, C, D, E, and F were collected in August 2016 and lastly, datasets G, H and I were collected in November 2016. Worth noting is the missing of 2015 data for Kenya and Uganda which was unrecoverable given the resources that were present at the time of data collection.

[4]https://cve.mitre.org/cgi-bin/cvename.cgi?name=cve-2015-0204
[5]https://cve.mitre.org/cgi-bin/cvename.cgi?name=cve-2014-0160

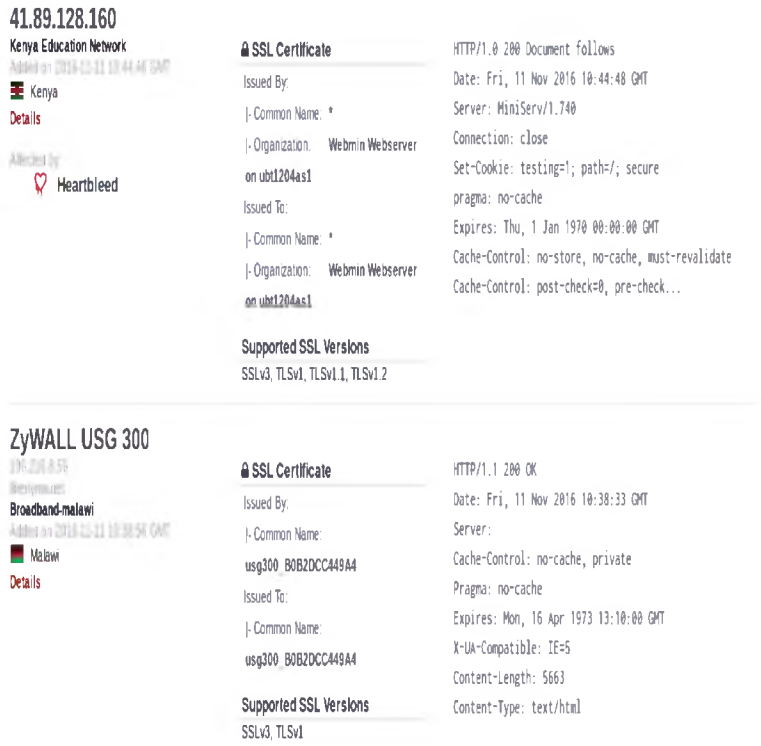


Figure 4.2: Sample Record of Shodan Output

Table 4.5: Shodan Default Password Dataset

Country	# of Records
Kenya	255
Tanzania	42
Malawi	11
Uganda	93
Total	402

4.6 Data Cleaning

Data processing was performed using a combination of common Unix text processing tools called *awk*, and a series of custom Python scripts, which also performed the analysis and plotting of data. Data collected from SHODAN does not come in a format legible by the human eye. Listing 4.1 shows a single row from SHODAN in *JSON* format before it was processed (cleaned) for use. The majority of the input data was JavaScript Object Notation (JSON) formatted data files, as obtained from SHODAN. The files had some inconsistencies making it impossible to be parsed directly. To make it readable one option was to read it line by line, get the string and use *json.dumps ('text')*. Alternatively turning the whole *.json* file into a list.

Table 4.6: Shodan Vulnerability Dataset

	Country				
Vulnerability	**Kenya**	**Tanzania**	**Uganda**	**Malawi**	**Total Records**
CVE-2015-0204	233	88	38	22	381
CVE-2014-0160	67	14	25	4	110
Total	300	102	63	26	491

Table 4.7: SHODAN Dataset Overview

	Time-frame		
Data set name	**(Month/Year)**	**Description**	**# of Records**
A	06/2015	Malawi_2015	3,836
B	06/2015	Tanzania_2015	15,513
C	08/2016	Kenya_2016	69,075
D	08/2016	Malawi_2016	6,632
E	08/2016	Tanzania_2016	19,778
F	08/2016	Uganda_2016	12,432
G	11/2016	CVE-1:Heartbleed Vulnerability	110
H	11/2016	CVE-2: FREAK Vulnerability	381
I	11/2016	Default Password	402
	Total		121,125

4.7 Data Characteristics

The study primarily focused on Internet Service Providers (ISP) that are providing the organization with the IP space to access the Internet, the type of device that accessed the Internet in order for SHODAN to acquire it's banner details. Such devises include but not limited to routers, web cameras, and switches. The research also looked at common ports that SHODAN used to grab the banners in Kenya, Uganda, Malawi, and Tanzania. Furthermore, it looked at the versions of OpenSSL that are used in the four countries under study i.e general purpose cryptography library that provides an open source implementation of the Secure Sockets Layer (SSL) and Transport Layer Security (TLS) protocols.

Lastly, the study looked at the products from various vendors that SHODAN managed to get hold of. These data categories were defined by SHODAN itself, this study just used them as they are i.e. they are known fields in SHODAN. Awk scripts were used to identify these data characteristics and a python code that was used for data extraction is attached to the appendix which shows the data characteristics that were extracted.

4.8 Summary

Chapter 4 defined data that was used in this study. All three categories of data sets used in this study (IBR data set, SHODAN 2015 data, and SHODAN 2016 data sets) are shown in this chapter. It defined what OSINT is in Section 4.1 followed by the legal aspects associated with using it and justification of why OSINT should be used to carry out studies in all fields. It also accommodated the shortcoming that comes with using it. IBR data was explained from how it was collected in Figure 4.1; what constituted it and the number of packets under study. This is the data that is used to have an understanding of information security in Chapter 5. Tables 4.1, 4.2 and 4.3 showed summaries of when the data was collected but also a complete data set overview. These tables are expanded further to give context in Chapter 5.

After explaining the characteristics of IBR data set, the study focused on SHODAN data sets, starting with the data collected in 2015 and the sample queries that were run to acquire data used in Chapter 6 and 7. This was explained in Section 4.5 together with naming the data sets based on the time of collection, the area from which the data was collected and the content of the data. This is summarised in Table 4.7. In the data sets collected in 2015, the focus was on three variables; ISPs, device types detected and ports. However, the scope expanded further when processing the 2016 SHODAN data set. SHODAN data was further broken down into eight different data sets indicating which country it belonged to but also the time it was collected. This is shown in Table 4.7 in Section 4.5. Tables 4.5 and 4.6 shows a summary of the data that was collected in November which is different from the data collected in 2015 and 2016 because of the content which it has. This data is labeled G,H, and I. More details on dataset I have been given in Chapter 6 while Chapter 7 has focuses sorely on dataset G and H. Sections 4.6 and 4.7 explained the data processing and data characteristics found in the SHODAN data.

Chapter 5

Information Security Perspective from Internet Background Radiation

This chapter introduces the reader to Internet Background Radiation (IBR) and explores IBR components that may be of great interest to security researchers regarding evidence of scanning activities, worms and misconfigurations (Moore, Shannon *et al.*, 2002; Yates, 2014). It goes on further to look at traffic composition based on the top IP source addresses found in the data collected using network telescope sensors. This study also investigates source address distribution across the IPv4 address space in /8, /16 and /24 network aggregates (Barford, Nowak, Willett, and Yegneswaran, 2006). In addition to this, it explores traffic composition based on TCP ports. This work also looks at port usage and why certain ports and services are targeted more than others.

The remainder of the chapter is structured as follows: Section 5.1 provides a brief introduction to IBR and its related work in the area. Section 5.2 explains the data analysis that was used to come up with the intended results which is followed by Section 5.3 which explains the ports that are being targeted in the dark net. Section 5.4 explains the significance of open ports. From ports, the study shifted its attention to source IP addresses in the countries under study. This is explained from Sections 5.5 through 5.9 which is followed by Section 5.10 which explains reused IP addresses in these four countries. Problems associated with geolocation are explained in Section 5.11 and the significance of unique IP addresses is given in Section 5.12. Recommendations are made in Section 5.13 and the chapter closes with a recap of what has been gathered from this chapter in Section 5.14

5.1 Related Work

IBR consists of non-productive data packets on the Internet (Cooke, Bailey, Mao, Watson, Jahanian, and McPherson, 2004) which are addressed to unused IP addresses or ports where there is no network device set up to receive them and often times shows evidence of either malicious activity or misconfiguration be it temporal or permanent misconfigurations (Pang *et al.*, 2004; Shannon and Moore, 2004). Given the fact that there are no legitimate hosts in an unused address block, also referred to as dark-space (Polakis, Kontaxis, Ioannidis, and Markatos, 2011; Irwin, 2012), then traffic must be the result of misconfiguration, backscatter from spoofed source addresses or scanning from worms and another probing (Cooke *et al.*, 2004; Wustrow *et al.*, 2010), it therefore, makes IBR an effective method for analysing and quantifying Internet security phenomenon.

The use of network telescopes, as a means of collecting IBR and analyzing network telescope traffic, has been used by security experts to understand the evolution of network threats and various potential malicious activity (Moore, Shannon, Voelker, and Savage, 2004; Irwin, 2011). This is done in order to give them an upper hand to develop applications and software that can cope with this change.

All the analysis shown in this chapter was based on datasets 146/8, 155/8 and 196/8 collected from three of Rhodes network telescopes. These data sets are shown in Table 4.3 and later on expanded to Table 4.3 and Table 4.4. The total number of packets was 258,156 of which Kenya had the majority. As explained in Section 4.3 only TCP packets were considered due to the absence or a small number of UDP and ICMP packets present in Uganda and Malawi respectively. All of the packets presented minus spoofing (which this research did not attempt to verify) is outbound traffic from EAC.

5.2 Data Analysis and Discussion

Table 5.1 is a sample representation of TCP destination port by country and percent composition of each port for each year. Table 5.1 shows top four ports in EAC throughout the study period of 72 months. Looking at Table 5.1, port 445 registered the highest traffic than any other port from 2009 to 2013. This is the case for Kenya, Uganda, and Tanzania. For Malawi, port 445 only registered more traffic in 2011,2012 and 2014. In most of these cases, port 445 registered more than half of the total traffic.

Table 5.1: Top TCP Destination Ports from EAC

		Kenya		Uganda		Tanzania		Malawi	
Year	**Rank**	**Ports**	**%**	**Ports**	**%**	**Ports**	**%**	**Ports**	**%**
2009	1	445	55.51	445	52.45	445	97.17	12203	22.16
	2	25	38.51	135	45.57	1433	1.96	2967	21.94
	3	135	3.78	80	1.35	5207	0.32	2968	21.56
	4	23	1.57	22	0.17	23	0.11	5900	21.56
2010	1	445	55.58	445	73.07	445	22.50	12203	19.32
	2	135	31.37	135	22.16	22	20.91	5900	18.83
	3	22	4.88	25	3.94	1433	2.93	2968	18.60
	4	21	2.12	22	0.12	139	1.29	2967	18.15
2011	1	445	52.11	445	90.97	445	74.15	445	76.19
	2	22	28.07	1433	5.63	22	12.02	3389	19.05
	3	139	10.31	80	3.11	443	6.04	3306	4.76
	4	1433	4.60	19216	0.11	1433	1.20	-	-
2012	1	445	49.68	445	76.08	445	62.60	445	42.15
	2	3389	21.18	22	18.37	22	16.90	3389	27.40
	3	22	9.03	23	5.11	23	6.45	22	10.96
	4	5900	8.63	3389	0.86	3389	5.95	3306	1.37
2013	1	445	60.73	445	71.14	445	40.61	80	43.97
	2	80	17.61	25	13.52	3389	18.28	445	43.10
	3	8080	10.96	3389	8.31	135	13.74	3389	8.97
	4	22	8.19	5900	4.39	80	11.21	25	0.52
2014	1	23	55.57	3303	34.63	22	40.99	445	91.60
	2	22	15.64	445	23.35	3389	18.67	3389	1.91
	3	445	11.45	7778	10.78	445	18.29	-	-
	4	3389	8.99	3389	10.23	80	12.50	-	-
2015	1	22	27.15	22	63.49	22	50.82	139	76.16
	2	80	23.46	445	16.18	445	26.53	445	23.18
	3	445	20.07	3389	12.45	80	12.45	-	-
	4	5916	17.63	23	6.22	8080	6.21	-	-

For 2014 and 2015 the traffic registered passing through port 445 reduced, instead port 22 and port 23 became more dominant especially in 2014 for Kenya (port 23 had 55.57% of the traffic), 2015 for Tanzania (port 22 registered 50.82% of the total traffic) and Uganda (port 22 registered 63.49% of the total traffic). As it can be seen from Table 5.1 the ports are arranged in order of percentage composition, with the highest being ranked position one.

After port 445 there was port 22 which was more dominant in 2014 and 2015 yet it seemed to have appeared throughout the study period registering its highest traffic in 2015 (63.49% of the total traffic in Uganda) and its lowest in 2009 in Uganda as well. Looking at Tanzania it was also easy to pick that port 1433/TCP was somewhat different from the other countries, appearing in 2009, 2010 and 2011. It also registered traffic in Uganda in 2011 and was second to port 445/TCP. TCP port 1433 offers Microsoft SQL Server services assigned by the Internet Assigned Number Authority (IANA) and was also a default port for SQL Server (Otey, 2014) i.e client systems use 1433/TCP to connect to the database engine. If the port is left insecure and not properly patched it becomes vulnerable to buffer overflow attacks (Speedguide, 2016a).

Malawi had unusual patterns compared to the other three countries in that other than port 445 being dominant in packet registration in 2011,2012 and 2014, other ports that were not appearing in the other three countries have high occurrence. For example, port 12203 which happens to be a UDP port was dominant in 2009 and 2010 respectively. Port 12203 is used for games when it uses UDP while it is unassigned under TCP (Speedguide, 2016b). There was also port 80/TCP which appeared at least once in each of the four countries often used by application protocols like the HyperText Transfer Protocol (HTTP) or on Web servers who opened this port solely for the purpose of listening to web traffic (Speedguide, 2016c). The alternative port used for web traffic is port 8080/TCP which also registered traffic in 2015 in Tanzania and 2014 in Kenya. In 2015, port 139/TCP registered traffic which was three times more than port 445. Port 139 is responsible for NetBIOS session services such as file sharing and printing occur here (Speedguide, 2016d). Once open it becomes prone to leaking system's information because it accesses the Internet thus making any poorly configured machine a risk and can only be resolved by disabling file and print sharing services or blocking the port entirely.

Table 5.2: Top 6 Destination Ports by Year

	TCP Port number and % Composition					
Year	**445**	**135**	**25**	**22**	**23**	**3389**
2009	58.44	14.93	24.12	0.67	0.99	0.85
2010	65.08	27.13	2.23	5.17	-	0.39
2011	75.91	0.34	0.19	23.13	-	0.43
2012	66.80	-	1.24	17.55	9.92	4.48
2013	68.75	7.24	0.01	8.38	-	15.62
2014	18.67	1.40	1.20	29.81	44.83	4.09
2015	33.52	-	-	55.65	6.69	4.14

5.3 Targeted Ports

This research looked at the ports that were being targeted and having done a geopolitical analysis, this section outlines the top six destination ports that were being targeted by the systems in the four countries together with their count. Previous studies and discussions in the security community have shown that port scans are a precursor to an attack (Spitzner, 2003; Provos *et al.*, 2004). It is for this reason that this study took interest in ports as one of its variables.

The count of ports covers all the 80 months period from August 2005 to June 2015. To come up with Table 5.2, ports with the highest count throughout the years were selected first and then split over the years. This was subsequently followed by computing their percent composition as this is more representative than their counts. The TCP ports that had the highest count throughout the 72 months period were ports 22, 23, 25, 135, 445 and 3389. In this section, this study looked at the services offered by each of these ports and vulnerabilities associated with them. In certain cases, solutions for the vulnerabilities was provided.

5.3.1 Port 445/TCP

Except for 2014 and 2015 port 445 had the highest composition in the traffic than any other port. What was more interesting was that during this same period, it was when port 80 and 8080 (not shown Table 5.2) began to appear in the network traffic. Services that run on port 445 is direct TCP/IP MS Networking access without the need for a NetBIOS layer (Mitre, 2010a). It is also responsible for file sharing in Windows and Microsoft-DS Active Directory (Porras, Saidi, and Yegneswaran, 2009). Having port 445 open, systems

and applications become vulnerable to certain worms such as W32.Deloader, IraqiWorm, W32.HLLW.Moega (Hayashi, 2004) and the Conficker worm (Porras *et al.*, 2009; Irwin, 2011).

Other worms that scan and capitalize on the vulnerabilities of port 445 during the period of observation are Korgo (Magee, 2007) and Sasser (Shannon, 2007). These worms are multitasking in nature and work by first exploiting the vulnerability and gaining a foothold in a host and then later move to self-propagation on that infected machine (Gates, Collins, Duggan, Kompanek, and Thomas, 2004; Irwin, 2012). Thus knowing that infected machines scan for vulnerabilities of port 445 (Gates *et al.*, 2004), as machines perform scanning for this port, there will be a large number of unique IP addresses contacted by a single source. It is therefore of no surprise that given this kind of worm behavior this might be the reason why there were more packets associated with port 445 in all the three datasets.

5.3.2 Port 135/TCP

Another port of consideration that appeared in the IBR dataset was port 135. This port is responsible for Remote Procedure Call (RPC) which involves two processes: the process doing the call and the process created to service the call (Schultz, 2004). Basically, it is a client-server set up meaning that it is possible to have different virtual or physical machines (Srinivasan, 1995). This is where port 135 comes into play to ensure that there is communication between the client and the server. The vulnerabilities associated with the port include buffer overflow condition in rpcss.exe (Microsoft, 2003), RPC Distributed Component Object Model (DCOM) vulnerability which occurs where a DCOM enabled service is listening (Kristensen, 2003). DCOM interface with RPC opens TCP port 135 to receive client requests for enabling DCOM objects.

These vulnerabilities also have worms associated with them. These include SBlaster worm, Bobax worm and Welchia (Nachia) worm that distorted the Internet in 2003 (Schultz, 2003). Traffic targeting port 135 was not as high as that of port 445 with no packets received in 2012 and 2015 and with almost negligible incidents in 2011 and 2014. This could be the result of a number of reasons: blocking ephemeral that RPC uses, blocking of all incoming port 135 traffic, host-based filtering, disabling RPC (Srinivasan, 1995; Schultz, 2004), patching done by the targeted organizations but also patching and security updates all of which seem to be possible solutions to such vulnerabilities and worm infections.

5.3.3 Port 25/TCP

The study also looked at port 25 which is responsible for all communication in regards to Simple Mail Transfer Protocol (SMTP) sessions. It allows an SMTP server and client to use Transport Layer Security (TLS) to provide private, authenticated communication over the Internet (Hoffman, 2002). Often times in the process of communication untrusted routers are involved which may give room to a third party monitor such transmission posing a security risk (Hoffman, 2002). Due to this set up a MiTM attack has the likelihood to occur either by the "250 STARTTLS" response from the server or allow the server to announce its STARTTLS capability for example (Hoffman, 2002).

Among the vulnerabilities that exploit port 25 include NetSPA which predicates its estimates on a network model based on vulnerability scans (Ingols, Chu, Lippmann, Webster, and Boyer, 2009). Many worms contain their own SMTP engine (Desouza, 2005) who use it to propagate by mass-mailing the payload. According to speed guide database integer overflow in Apple Safari in CVE-2010-1099 (Mitre, 2010b), Arora in CVE-2010-1100 (Mitre, 2010c), Alexander Clauss iCab in CVE-2010-1101 (Mitre, 2010d), OmniWeb in CVE-2010-1102 (Mitre, 2010e) and Stainless in CVE-2010-1103 (Mitre, 2010f) allows remote attackers to bypass intended port restrictions on outbound TCP connections using the same TCP port, 25. Among the long list of trojan horses that exploit this port include Ajan, Antigen, Barok, Email Password Sender - EPS, EPS II, Moscow Email trojan, Naebi, NewApt worm and ProMail trojan (Speedguide, 2015). The packets received by the network telescope showed that 2009 showed more packets than all the other years that followed.

With each passing year, the traffic constituted of less than 2% except for 2010 which registered 2.23% of the total packets. This could be as a result of following proper security procedures like patching and disabling of unnecessary services on this port. A good example of such desirable results is that of 2015 where no packets were registered under port 25. The same result almost happened again in 2011 and 2013 where only 0.19% and 0.01% of the total traffic observed were registered for each year respectively.

5.3.4 Port 22/TCP

To ensure encrypted and secure communication of any service that uses command line access then port 22 is the route to go as it is responsible for secure shell communication (Desouza, 2005). Due to the nature of the service it offers, it makes it susceptible to various

forms of remote attacks like DoS (crash) via an SSH2_MSG_NEWKEYS packet to TCP port 22, which triggers a NULL pointer dereference as shown in CVE-2008-0852 (Mitre, 2008). Not only that but also unauthenticated remote attacker with network access to port 22 can tunnel random TCP traffic to other hosts on the network via Ruckus devices (Speedguide, 2016e). The worst scenario is shown by CVE- 2012-4702 (Mitre, 2012a) where a remote attacker could gain root privileges on 360 systems Maxx via an SSH session.

Other than traffic recorded in 2012 (17.55%) and 2013 (8.38%) port 22 showed a consistent rise in the traffic targeting it. With 2015 showing the highest traffic ever recorded for port 22 (55.67%); almost double as much as that of its preceding year (29.81%), surpassing that of port 445 (33.52%). This increase may be attributed to an increase in activities that required the use of SSH than its preceding years.

5.3.5 Port 23/TCP

Despite being one of the oldest Internet protocols, with most popular programs for remote access to Unix machines, telnet, which is also offered as a service by port 23, has numerous security vulnerabilities attached to it (Postel and Reynold, 1983). Among the vulnerabilities include allowing remote attackers to cause DoS (device restart) via a crafted packet (Mitre, 2010g), buffer overflow in the Remote command server (Rcmd.bat) in tiny TCP/IP server allowing remote attackers to cause a DoS (Mitre, 2012b). Other trojans like DM worm, Aphex's Remote Packet Sniffer, AutoSpY, ButtMan, Fire HacKer, My Very Own trojan, Pest, Mirai botnet IoT and RTB 666 (Desouza, 2005; Dobbins, 2016) also make use of this vulnerability.

Port 23 in reference to Table 5.2 showed a lot of incidents with zero traffic. These include 2010, 2011 and 2013. Its highest traffic was recorded in 2014 (44.83%) making it the highest port targeted in 2014. For the first three years i.e. 2009 - 2011, its packet composition was almost negligible (in reference to 2009) until 2012. It went to zero again in 2013 making its mark in 2014 and went down again 2015. Despite the fact that it has three incidents that have zero registration of packets by the sensors, the number of packets that were recorded in those other four years surpasses that of the other ports that are not included on this list.

5.3.6 Port 3389/TCP

Services that require Remote Desktop Protocol (RDP) for accessing systems over the Internet, especially in server environments port 3389 is the channel to use (Desouza, 2005). However, this port has its vulnerabilities once its left open; Microsoft Security bulletin number MS02-051 and MS01-040 have more details on this. For instance, trojans that make use of this vulnerability include backdoor.Win32.Agent.cdm and TSPY_AGENT.ADDQ (MIcrosoft, 2012). This port is also vulnerable to DoS against Windows platforms as a remote attacker can quickly cause a server to reach full memory utilization by creating a large number of normal TCP connections to port 3389 (Mitre, 2012c; Hunter *et al.*, 2013).

There is also another bug that could allow a remote unauthenticated attacker to run arbitrary code on the affected system by sending a sequence of specially crafted RDP packets to port 3389/TCP which gives the RDP access to any object in memory even after it has been deleted (Mitre, 1999). Port 3389 showed the least traffic compared to any other ports that were in the top six. It never registered zero traffic but its numbers were consistently low throughout the data collection period. With its lowest composition in 2010 (0.39%) followed by 2011 (0.43%) then 2009 (0.85%).

5.4 Significance of Open Ports

Section 5.3 has looked at the various ports that SHODAN gets to pull banners from. More importantly, it has looked at various services that run on these ports in order to ensure communication between devices on any given network. It is worth noting that any of the given services that run on these port are supporting a specific program and as such, they possess information that could prove useful to an attacker. If a pseudo program has a vulnerability and runs on any of the open ports, then it can be attacked on the port it is assigned to. All that is remaining is for an attacker to map the program to the port it is listening to since it's not possible to attack a program on ports it is not listening to.

A newly installed operating system (OS) often has a number of things running automatically, some of which being services which facilitate connectivity on a network. Any network exploitable vulnerability in such a service running on open or insecure ports is a potential access door for an attacker. Blocking access to a given port can be done on the firewall and is considered more efficient than stopping an OS from running a given service at the occasion of a software update giving more access to attackers. So it is customary

to block all ports except those which are known to correspond to services which should be accessible worldwide (e.g. 80/TCP and 22/TCP, for Web and SSH, respectively) which should be kept secure and up to date.

5.5 Source IP Addresses

Having established that at times persistent request connections may be a sign of a widespread of an infection, like Conficker worm, from a unique source IP address (Irwin, 2012), a selection of the dominant top four source netblock address blocks were selected from the observed traffic for each country under study. An example of such computation is shown in Table 5.3 where a breakdown of top four source netblock addresses for each year i.e. from 2009 to 2015 is computed to show how much they contributed to the entire traffic. Then its packet composition in relation to the total number of packets received from various unique IP sources captured by the network telescope was computed. Note that the last octet of the IP addresses has been concealed (hence the word netblock) for security and privacy of the owners of the IP addresses and the IP addresses were collapsed to /24 for reporting but also provide the total count shown in Table 5.3. From here onwards the study will refer to these sources IP who have the identity of the last octet concealed as source netblock address.

In reference to Table 5.3, source netblock 196.201.141.X registered 28.60% of the total traffic in that year, in 2010 it was 196.201.208.X registering 9.88%. Due to the distribution of the traffic in 2010 one can easily notice that the distribution was even with top four only registering 27.29% of the total traffic meaning 72.71% was distributed among the other net-blocks.

A similar case was seen from 2011 to 2012 with the top netblock 41.139.193.X having registered 12.93% of the traffic and 41.139.255.X registered 18.40% respectively. In 2013 netblock 197.232.13.X registered 21.51% of the total traffic and 212.49.70.X registered 34.20% of the traffic. Unlike the situation in 2010 and 2011, 2015's top four net-blocks registered 76.0% of the total traffic with the top netblock 212.49.70.X registering 25.73% of the total traffic.

Table 5.4 shows Tanzania's top four net-blocks distributed by year in which they registered traffic but also the percent composition of the traffic they registered. Netblock

Table 5.3: Kenya's Top 4 Source /24 Net-blocks by Composition per Year

Year	Rank	Src Netblock	%	# of packets
2009	1	**196.201.141.X**	**28.60**	**8,142**
	2	196.201.148.X	16.06	4,570
	3	196.200.21.X	3.68	1,048
	4	196.201.208.X	3.44	980
2010	1	**196.201.208.X**	**9.88**	**2,330**
	3	196.202.212.X	6.30	1,800
	3	196.201.226.X	3.91	1,478
	4	62.24.111.X	3.29	1,004
2011	1	**41.139.193.X**	**12.93**	**2,077**
	2	196.202.196.X	12.33	972
	3	41.204.167.X	3.95	563
	4	212.49.95.X	6.78	595
2012	1	**41.139.255.X**	**18.40**	**1,531**
	2	196.202.214.X	13.56	1,128
	3	41.204.168.X	12.97	1,079
	4	217.199.145.X	11.50	957
2013	1	**197.232.13.X**	**21.51**	**2,468**
	2	196.201.211.X	17.85	3,072
	3	196.201.208.X	13.39	1,536
	4	196.202.202.X	4.66	599
2014	1	**41.222.15.X**	**24.30**	**10,178**
	2	212.49.70.X	14.69	6,140
	3	41.215.28.X	3.68	1,540
	4	197.211.12.X	2.38	995
2015	1	**212.49.70.X**	**25.73**	**2,913**
	2	41.203.214.X	20.54	2,823
	3	197.248.144.X	20.21	2,288
	4	197.232.26.X	10.02	1,136

Table 5.4: Tanzania's Top 4 Source /24 Net-blocks by Composition per Year

Year	Rank	Src Netblock	%	# of packets
2009	1	**196.44.161.X**	**23.94**	**896**
	2	196.46.129.X	5.64	291
	3	196.43.78.X	4.81	246
	4	196.41.45.X	4.30	239
2010	1	**41.223.4.X**	**26.12**	**991**
	2	196.43.67.X	13.44	539
	3	196.44.161.X	12.89	518
	4	196.45.146.X	8.56	335
2011	**1**	**41.220.180.X**	**11.78**	**499**
	2	196.43.84.X	7.32	400
	3	196.44.162.X	6.04	328
	4	196.1.53.X	5.97	256
2012	1	**41.77.228.X**	9.21	**744**
	2	196.45.156.X	9.18	741
	3	41.59.13.X	6.37	514
	4	196.44.173.X	6.64	512
2013	1	**196.45.144.X**	**23.21**	**2,257**
	2	41.59.7.X	13.22	1,286
	3	196.43.84.X	12.21	1,187
	4	196.41.43.X	7.90	768
2014	1	**196.46.100.X**	39.27	**7,781**
	2	196.41.43.X	15.91	3,153
	3	41.93.45.X	15.23	3,072
	4	41.59.61.X	4.35	861
2015	1	**196.46.100.X**	**45.74**	**2,694**
	2	196.41.50.X	22.94	1,353
	3	196.45.144.X	6.17	364
	4	196.41.47.X	5.16	304

Table 5.5: Malawi's Top 4 Source /24 Net-blocks by Composition per Year

Year	Rank	Src Netblock	%	# of packets
2009	1	**41.221.106.X**	**93.84**	**1,249**
	2	196.45.190.X	3.23	43
	3	41.221.96.X	0.90	14
	4	196.45.189.X	0.45	7
2010	1	**41.221.106.X**	**75.30**	**1,668**
	2	196.216.8.X	11.65	258
	3	41.221.99.X	9.48	210
	4	41.221.97.X	2.08	38
2011	1	**196.216.10.X**	**50.00**	**12**
	2	41.221.96.X	25.00	6
	3	41.77.13.X	12.50	3
	4	41.221.100.X	6.25	2
2012	1	**41.221.96.X**	68.81	**75**
	2	41.77.12.X	22.02	24
	3	41.78.57.X	9.17	10
2013	1	**41.77.13.X**	**43.74**	**255**
	2	41.221.97.X	36.21	210
	3	41.87.10.X	9.14	53
	4	105.234.255 .X	2.41	14
2014	1	41.221.97.X	**75.19**	**197**
	2	41.77.14.X	4.58	12
	3	41.221.108.X	1.91	5
2015	1	**41.221.103.X**	**90.91**	**150**
	2	41.77.14.X	4.84	9
	3	41.75.112.X	1.82	5

196.44.161.X registered 23.94% of the traffic in 2009 representing close to a quarter of the total traffic. For the other three in the top four, there seems to be a close range of the amount of traffic that passed through them when compared to 196.44.161.X. In 2010 netblock 41.223.4.X registered 26.12% of the total traffic. Unlike in 2009, 2011 and 2012, the net-blocks in 2010 registered more than half of the total traffic (61.01%) registered in Tanzania. The same thing happened in 2013 except this time the top four registered over 55% of the total traffic with the top netblock 196.45.144.X registering 23.21% of the total traffic followed by 13.22% from 41.59.7.X and 196.43.84.X with 12.21%. From 2013 up to 2015 the amount of traffic registered by the top IP addresses keeps on increasing, from 23.21% in 2013, then 39.27% in 2014 and 45.74% in 2015.

Malawi's net-blocks in Table 5.5 show two different patterns compared to the other three countries under study in that apart from 2011 and 2013 the rest of the top net-blocks recorded more than half of the total traffic. To begin with, in 2009 netblock

41.221.106.X registered 93.84%, the same netblock (41.221.106.X) registered 75.30% in 2010, 196.216.10.X registered 50.00% in 2011, 41.221.96.X registered 68.81% in 2012 and 41.221.103.X registered 90.91%. Secondly, all the top four net-blocks in Malawi registered more than two-thirds of the total traffic throughout the study period. This is not the case for Kenya, Tanzania, and Uganda where the top four failed to register more than half of the traffic in certain periods.

Apart from 2010 and 2015, Uganda showed moderate margins between the topmost net-block and the other three that follow as it is evident from Table 5.6. From this table, it can be seen that in 2010 the top netblock that registered the highest traffic is 196.0.26.X (69.85%), the second one registered 10.74% (i.e. six times more than 196.0.19.X). In 2011 the topmost netblock 196.0.4.X registered 28.43% follow by 11.50% recorded by IP address block 196.0.5.X. In 2015, netblock 154.0.130.X registered 63.79% of the total traffic. Most of the top net-blocks in Uganda registered at least half of the total traffic compared to the other net-blocks in the given study period.

Given the nature of net-blocks, it was not possible to aggregate them as we would with ports since each of them is unique to each country. As such we will look at four different cases: Kenya, Malawi, Uganda, and Tanzania. In each case study the results, represented in tables, will show the top source IP address blocks that sent more packets than the rest of each year of data collection. They will also be categorized according to the year in which addresses were used.

5.6 Kenya

Table 5.7 shows a breakdown of /24 top net-blocks and their ISP recorded in by the network telescope sensors in Kenya. It also shows the number of packets received by each of the net-blocks.

The tables further show that most of these unallocated net-blocks belonged to Safaricom Limited. In terms of geographical location, all these ISPs were located in Nairobi. 2009 registered more packets than any other year. From the original dataset containing all IP addresses it showed that some of the source net-blocks were used more than once. For example, 212.22.182.X belonging to Internet-Solutions-Ke as its ISP appearing both in 2009 and 2010 respectively. Of interest was the fact that it used a different port in 2010 than the one it used in 2009.

Table 5.6: Uganda's Top 4 Source /24 Net-blocks by Composition per Year

Year	Rank	Src Netblock	%	# of Packets
2009	1	**196.0.13.X**	**20.60**	**2,678**
	2	196.0.11.X	18.16	2,361
	3	196.0.5.X	6.80	881
	4	196.0.17.X	5.78	752
2010	1	**196.0.26.X**	**69.85**	**13,422**
	2	196.0.19.X	10.74	2,063
	3	196.0.12.X	4.01	760
	4	196.0.25.X	3.91	753
2011	1	**196.0.4.X**	**28.43**	**769**
	2	196.0.5.X	11.50	311
	3	81.199.21.X	5.62	152
	4	196.43.133.X	5.43	147
2012	1	**196.0.4.X**	33.81	**1,711**
	2	**196.0.41.X**	31.11	1,574
	3	196.0.7.X	10.49	531
	4	196.0.26.X	6.74	341
2013	1	**196.0.4.X**	**15.69**	**270**
	2	196.0.31.X	13.42	231
	3	41.223.85.X	12.78	220
	4	196.0.64.X	9.93	171
2014	1	**196.0.29.X**	**34.62**	**768**
	2	212.88.100.X	18.85	418
	3	196.0.26.X	10.78	239
	4	196.0.64.X	8.16	181
2015	1	**154.0.130.X**	63.79	**768**
	2	41.221.84.X	8.39	101
	3	196.0.64.X	6.06	73
	4	212.88.112.X	3.65	44

Table 5.7: Top Source /24 IP Net-blocks by Year in Kenya

Year	Src Netblock	ISP	%
2009	196.201.141.X	Iway-noc	28.60
2010	196.201.208.X	Safaricom Ltd	9.88
2011	41.139.193.X	Safaricom Ltd	12.93
2012	41.139.255.X	Broadband-adsl	18.40
2013	197.232.13.X	Safaricom Ltd	21.51
2014	212.49.70.X	Telkom Kenya Ltd	34.20
2015	197.248.144.X	One Communications Ltd	27.96

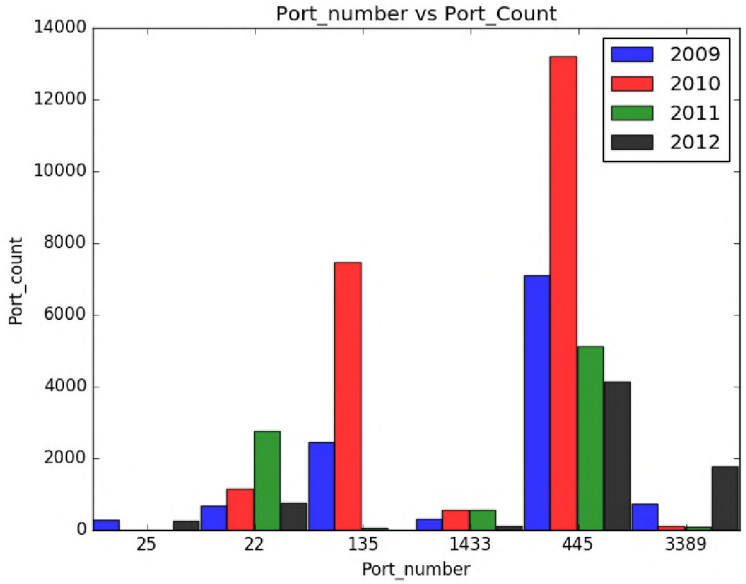


Figure 5.1: Kenya's Top Targeted Ports: 2009 - 2012

Having the same source IP may be an indication that the IP still served the interests of the malicious activities it was intended for as it can constantly be used without being attended to. But it also means that nothing has been done about it since it was detected. Other source net-blocks that were reused in other years include 212.49.70.X which appeared in 2014 and 2015 but also 212.49.95.X. In each case, the IP addresses used a different port.

Figure 5.1, 5.2 and 5.3 show the port distribution in Kenya that was detected by the telescopes at Rhodes University. The objective of these graphs was to show how ports were detected over time to see if there were changes to them i.e. an increase in a specific port or decrease or any form of pattern that cannot otherwise be picked at random. Figure 5.1 shows ports that were registered from 2009 to 2012. From the graph, it is clear that port 445 registered more traffic than any other throughout the study period with its highest record being in 2010. But looking at Figure 5.3 it shows that during the period of 2013 to 2015 it was port 23 that registered the highest record with over 20000 ports. Figure 5.2 is similar to 5.3 with the only difference being the elimination of port 23 from the equation. This was done because there was a huge margin between the port 23 and its closest competitor making the graph hide some of the essential details that could not otherwise be seen and render them insignificant.

The gaps seen in Figure 5.1 and 5.3 are a result of those ports registering zero traffic in those particular years or that their traffic was insignificant. For example port 25 in Figure 5.1, shows that in 2010 and 2011 it never registered any traffic. It also shows that

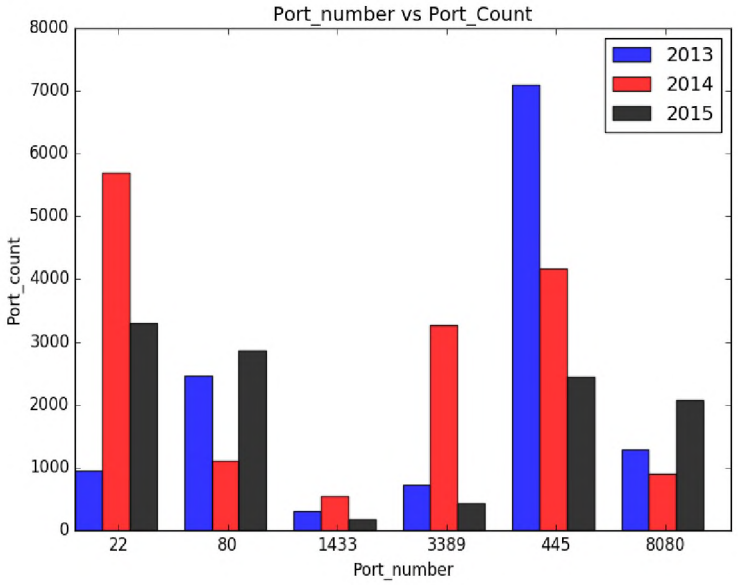


Figure 5.2: Kenya's Top Targeted Ports: 2013 - 2015

the highest traffic was registered in 2010 evidence from the count of port 135/TCP and 445/TCP respectively. Very little traffic was detected in 2012. From 2010 onwards there was a decline in the traffic registered by port 445/TCP. This trend continued from 2013 onwards up until 2015. While port 445 traffic is declining port 22/TCP shows an increase in traffic registered in 2009 to 2011 as seen in Figure 5.1. For the other ports, there is no consistent pattern with one point showing a decline in one year and the following year a hike, as seen with port 3389/TCP, 80/TCP and 8080.

5.7 Malawi

Table 5.8 shows a breakdown of top IP addresses and their Internet Service Provider (ISP) recorded from Malawi's outbound traffic by Rhodes network telescope. Malawi Telecommunications Ltd (MTL) is the ISP that that dominated in terms of source IP addresses that registered more traffic.

This dominance could also be attributed to the fact that MTL is one of the oldest (if not the oldest) ISPs in Malawi and is owned by the government. It also had source net-blocks appearing more than an allocated year of data collection. These include 41.221.96.X which were picked by the telescope from 2010 to 2013 and 41.221.97.X and which appeared in the traffic in 2013 and 2014 and each time it used a different port. Both of these belonged to MTL. Just like Kenya, all of these ISPs are located in one city, Blantyre.

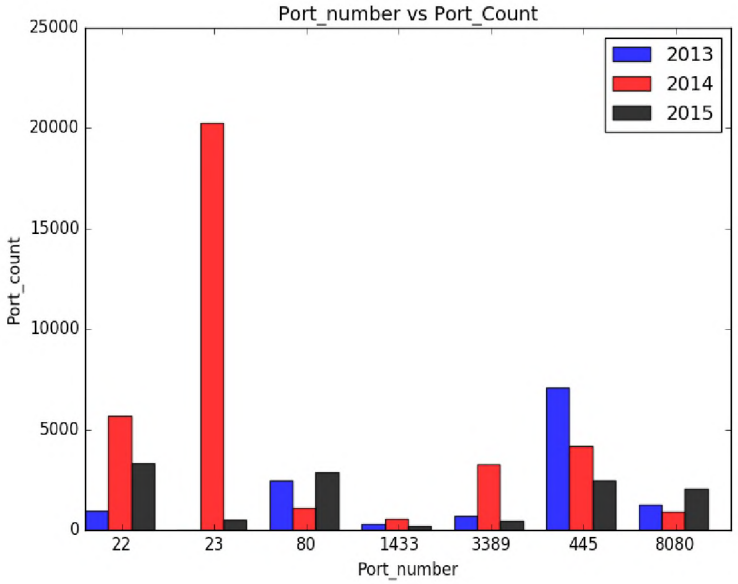


Figure 5.3: Kenya's Top Targeted Ports: 2013 - 2015 (With port 23/TCP)

Table 5.8: Top Source /24 Net-blocks by Year in Malawi

Year	Src Netblock	ISP	% Composition
2009	41.221.106.X	MTL	93.84
2010	41.221.106.X	MTL	75.30
2011	196.216.10.X	MTL	50.00
2012	41.221.96.X	MTL	55.83
2013	41.77.13.X	Globe	43.74
2014	41.221.108.X	MTL	41.03
2015	41.221.103.X	MTL	90.91

Figures 5.4 and 5.5 show the port distribution in Malawi covering the period of 2009 to 2015. Traffic registration in Malawi was not consistent in that, other than port 445 in the period of 2009 to 2012, no other port registered traffic consistently for three years. This is the case even for the years 2013 through 2015. Another inconsistency was the ports used evidently in Figure 5.4 and 5.5 where only ports 445 and 3389 were used after 2012. There was a change in ports usage where entirely new ports were brought into the picture. It can also be seen that there were fewer packets registered on port 25, 80, 22 and 139.

In the period 2009 to 2013 port 12203/TCP registered more packets than any other port in 2009 and 2010 there was an equal amount of traffic passing through port 12203 and 2967. Looking at Figure 5.4 it can be seen that there was almost an equal distribution of packets in the ports that registered traffic. This was the case for ports 12203/TCP,

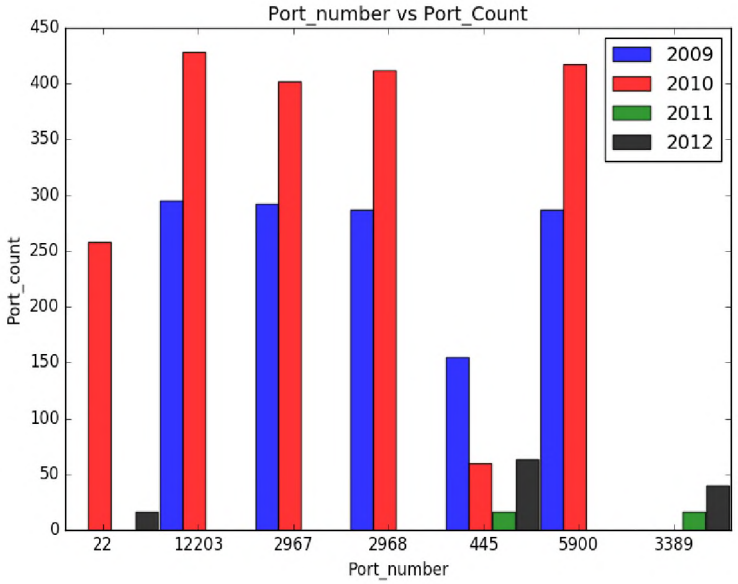


Figure 5.4: Malawi's Top Targeted Ports: 2009 - 2012

2967/TCP, 2968/TCP and 5900/TCP for both 2009 and 2010. Based on these two graphs, it is evident that Malawi registered very little IBR traffic in 2012. Worth noting is the fact that the ports that registered traffic were unique only to Malawi i.e. ports 12203, 2967, 2968 and 5900. In the other three countries, they may have registered traffic but it was not quite significant as it was in the case with Malawi.

5.8 Tanzania

Table 5.9 shows the ISPs, number of packets and source IP addresses that were recorded from Tanzania's outbound traffic by Rhodes network telescope. With all of the ISPs listed in Table 5.9, situated in Dar es Salaam. University of Dar es Salaam's source IP addresses recorded the highest number of packets in 2009 and 2012. Despite this being the case, large amounts of traffic were observed in 2014 with the IP address belonging to Tanzania Telecommunications Limited (TTCL).

Like the previous two countries, Tanzania had net-blocks that appeared more than once: 196.43.78.X was registered in 2009 and 2010, 196.43.84.X was registered in 2013 and 2014, 196.46.100.X was registered in 2014 and 2015. More evidence that unallocated IP addresses can be reused for further malicious activity if the addresses are not properly secured.

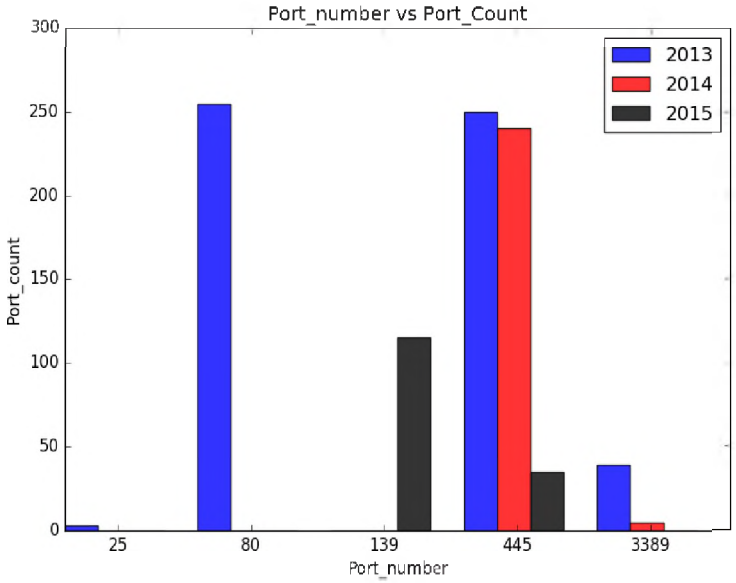


Figure 5.5: Malawi's Top Targeted Ports: 2013 - 2015

Table 5.9: Top Source /24 IP Net-blocks by Year in Tanzania

Year	Src Netblock	ISP	%
2009	196.44.161.X	University of Dar es Salaam	23.94
2010	41.223.4.X	Vodacom	26.17
2011	41.220.180.X	Africa Online (T) Ltd	11.78
2012	196.44.173.X	University of Dar es Salaam	6.78
2013	196.45.144.X	Cats-Net Limited	33.81
2014	196.46.100.X	TTCL	44.26
2015	196.46.100.X	Startel Tanzania Ltd	48.04

Figures 5.6 and 5.7 shows ports that were targeted in Tanzania from 2009 to 2015. More evident is the pattern of port 445 and port 135 in Figure 5.7. These two show a continual decline from 2013 onwards with each registering it's highest packets count in 2013 and lowest in 2015. Port 135 never registered any traffic in 2015 however, port 445 was dominant in the period 2009 to 2012 while port 22 registered the highest possible number of packets in Tanzania throughout the study period. Another interesting item of note was port 23 which never registered any traffic until 2012 and never appeared again. In 2009 there was no traffic for port 22 and 3389 and that was the case as well for port 1433 in 2012. Figure 5.6 also shows that from 2010 to 2012 port 445 began registering more traffic than any other port available.

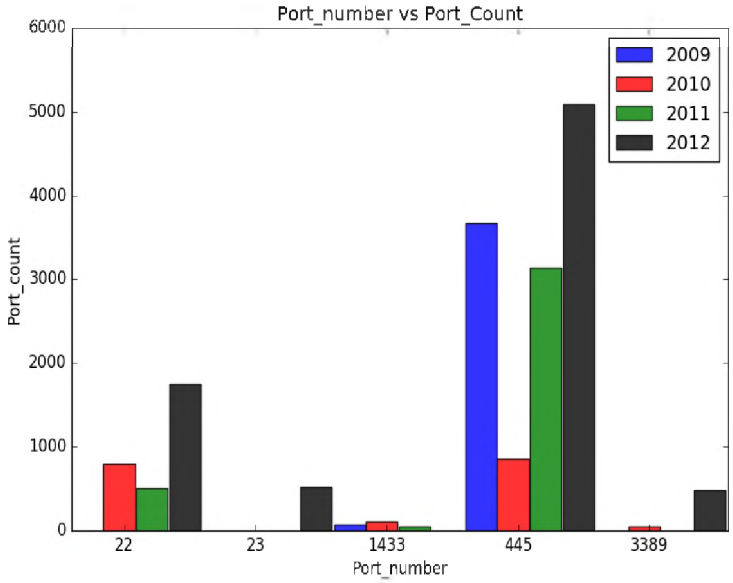


Figure 5.6: Tanzania's Top Target Ports: 2009 - 2012

Table 5.10: Top Source /24 Net-blocks by Year in Uganda

Year	Src Netblock	ISP	% Composition
2009	196.0.13. X	Uganda Telecom Ltd	19.04
2010	196.0.26.X	Uganda Telecom Ltd	70.73
2011	196.0.4.X	Uganda Telecom Ltd	24.80
2012	196.0.41.X	Uganda Telecom Ltd	42.68
2013	196.0.31.X	Uganda Telecom Ltd	14.49
2014	196.0.29.X	Uganda Telecom Ltd	35.15
2015	154.0.130.X	RII	74.06

5.9 Uganda

The last case study is that of Uganda. Table 5.10 gives a summary of its findings. Except for netblock 154.0.130.X which belonged to Roke Investments International (RII), all of the ports that registered the highest level of traffic in their respective years belonged to Ugandan Telecom Limited. Like the other three countries in this study, all the ISPs were situated in Kampala. 2010 registered the highest number of packets (13,370) not only in Uganda but for the entire data set in all four countries. Netblock 196.0.13.X registered packets for both 2009 and 2010, with its record of 2009 being the highest record for that year while 41.221.84.X and 41.190.212.X also registered packets in two consecutive years (2014 and 2015).

Apart from Malawi, Uganda was the other country under study that registered little

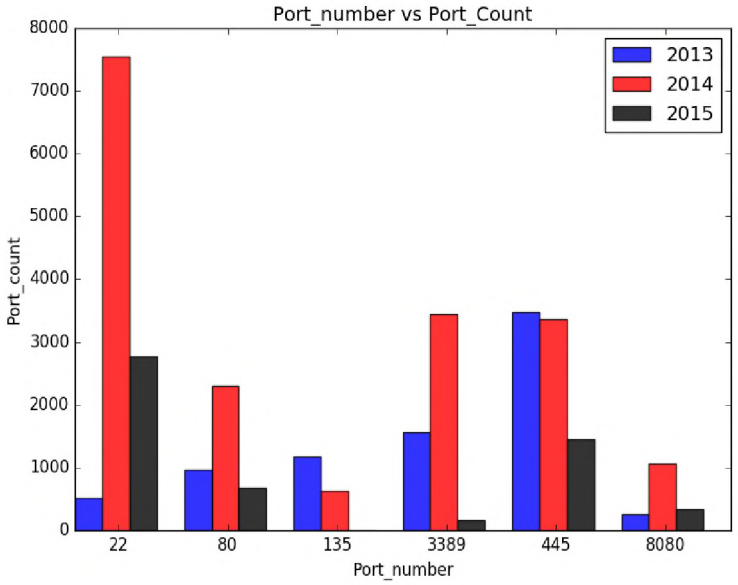


Figure 5.7: Tanzania's Top Target Ports: 2013 - 2015

traffic. Figures 5.8 and 5.9 shows that port 22/TCP, 23/TCP and 25/TCP registered the least of the traffic with port 22 in the period 2009 to 2012 registering no traffic until 2012. While port 25 only registered traffic in 2010 and 2013. Port 445 registered more traffic than any other port throughout the study period. From 2012 on wards port 445/TCP showed a continual decline in the number of packets that it registered. The same goes for port 135 which registered more packets in 2009 but after that it was a continual decline and after 2010, it never appeared again among the exploitable ports. As port 445/TCP and 135/TCP were showing a decline in traffic registered, port 23/TCP which was not present until 2014, was increasing its registration of packets with its highest being recored in 2015. From Figure 5.9, only port 25/TCP seemed to not have registered any traffic in 2015

5.10 Reused IP Addresses

Often times cybersecurity attackers tend to reuse their resources and tools as a way of verifying if some changes have been made to the system that they attacked before. The reuse of IP addresses found in IBR data is one of those techniques that attackers do by randomly probing IP addresses they have used before. At times it becomes possible that they have been holding on to it until it is allocated or it could be that the company that owns the range of IP addresses under probe has not done anything to secure its network.

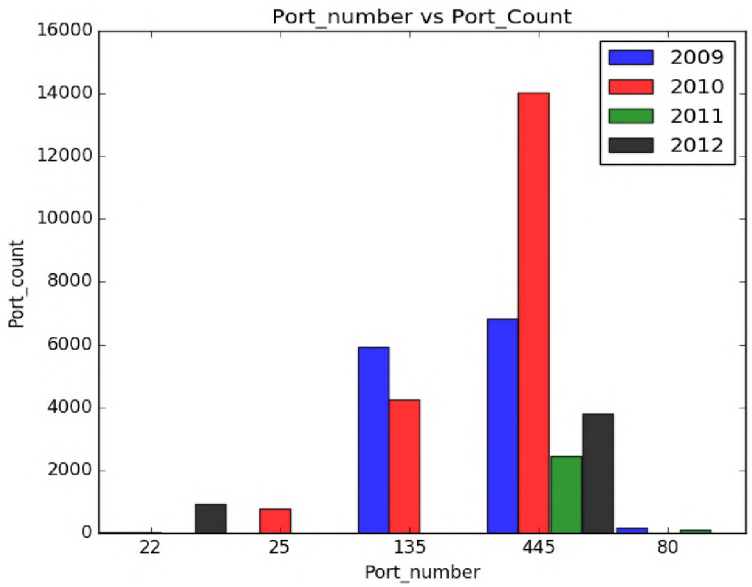


Figure 5.8: Uganda's Top Target Ports: 2009 - 2012

Either way none of the aforementioned above is good for the owner. Repeated appearance of certain IP addresses indicates that the given IP address is under close monitoring by someone else other than the owner and should this continue, the network hosting the IP address is placed in great danger.

Table 5.11 lists all net-blocks that were reused at least twice during the 72 month study period. Some of them were used in consecutive years while other had gap years in between them. Table 5.11 is presented in a way that for each year that the IP address was used, it recorded the column named after its country of origin. For example, netblock 196.201.212.X found in Kenya was used in 2009 and reused 2010 and 2011 thus if we go back to the table we expect to find this netblock in Table 5.11 in all the years that have been mentioned here. Looking at Kenya's net-blocks it is seen that IP netblock 196.201.208.X was used in 2009 and reused 2010 and 2013, netblock 196.200.26.X was used in 2009 and reused in 2010. Netblock 196.202.202.X was used in 2012 and reused in 2013 and 2015. Most of the net-blocks that were used in 2009 and 2010 were used again in later years with 2012 being the least victim of IP address reuse.

Other than 2011, Uganda's IP address reuse had been quite high. In fact, it was the one leading in IP address reuse compared to the other three countries in EAC. Netblock 196.0.13.X was used in 2009 and reused in 2010, 196.0.25.X was used in 2009 and reused in 2010 and 2012, 41.190.212.X was used in 2013 and reused in 2014 and 2015. Netblock 196.0.17.X was used in 2009 and reused in 2012, 2013 and 2014, it is the most reused

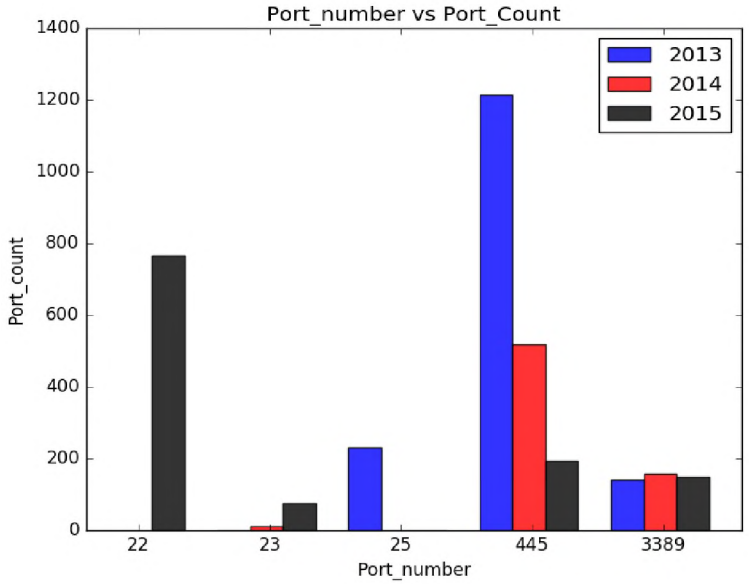


Figure 5.9: Uganda's Top Target Ports: 2013 - 2015

netblock in Uganda appearing four times just like IP netblock 196.44.162.X in Tanzania used in 2010 and reused in 2011, 2012 and 2013. In Tanzania IP netblock 41.77.228.X was used in 2011 and reused in 2014 and 2015, IP netblock 196.46.122.X was used in 2009 and reused in 2012. In Malawi, 41.77.228.X, appearing five times, was used more than any netblock throughout the study period, first in 2009 then reused in 2010, 2011, 2012, 2013 and 2014. Then there is IP netblock 41.221.97.X used in 2010 and reused in 2013 and 2014.

5.11 Problems with Geolocation

Using the same source IP addresses this research investigated source address distribution across the IPv4 address space to identify the geolocation of the IP addresses (Barford *et al.*, 2006), i.e. are the addresses tightly located in the same city or dispersed across the countries of origin. It also looked at the service providers of these IP addresses. Based on the data that was used in this study one thing stood out; the data was not properly represented. This is the case because from all the four countries in this study all of the IP addresses seemed to have come from one city, i.e for Malawi it was Blantyre; for Tanzania it was Dar es Salaam; for Uganda it was Kampala and Kenya it was Nairobi. This portrays a picture that IBR only occurs in four cities which is not entirely true. The data could have been more representable if the IP addresses were distributed across each

Table 5.11: Reused /24 Net-blocks

	Kenya	Uganda	Tanzania	Malawi
Year	**IP address**	**IP address**	**IP address**	**IP address**
2009	**196.201.212.X**	196.0.13.X	**196.46.122.X**	41.221.106.X
	196.201.208.X	**196.0.25.X**	196.44.161.X	**41.221.96.X**
	196.200.26.X	196.0.5.X	196.41.36.X	-
	212.22.182.X	196.0.18.X	196.43.78.X	-
	196.200.21.X	**196.0.17.X**	196.41.45.X	-
	212.49.95.X	196.0.12.X	-	-
2010	**196.202.212.X**	196.0.13.X	196.44.161.X	41.221.106.X
	196.201.208.X	**196.0.25.X**	196.43.78.X	**41.221.96.X**
	196.200.26.X	196.0.18.X	196.43.84.X	**41.221.97.X**
	212.22.182.X	196.0.12.X	**196.44.162.X**	-
	196.200.21.X	41.221.92.X	196.45.144.X	-
	212.49.95.X	196.0.7.X	-	-
2011	212.49.95.X	196.0.5.X	196.41.36.X	**41.221.96.X**
	41.204.167.X	212.88.119.X	196.41.45.X	196.216.10.X
	196.202.212.X	196.0.4.X	196.43.84.X	41.221.100.X
	-	196.0.3.X	**196.44.162.X**	41.77.13.X
	-	-	196.45.158.X	-
	-	-	196.46.120.X	-
2012	**196.202.202.X**	**196.0.25.X**	**196.46.122.X**	**41.221.96.X**
	-	**196.0.17.X**	**196.44.162.X**	196.216.10.X
	-	196.0.7.X	196.45.158.X	41.77.12.X
	-	196.0.26.X	196.46.120.X	-
	-	196.0.4.X	41.59.13.X	-
	-	196.0.3.X	-	-
2013	**196.201.208.X**	**196.0.17.X**	196.43.84.X	**41.221.96.X**
	212.49.95.X	41.221.92.X	**196.44.162.X**	**41.221.97.X**
	196.201.211.X	196.0.4.X	196.45.144.X	41.221.100.X
	196.202.202.X	41.190.212.X	196.41.43.X	41.77.13.X
	-	196.0.64.X	196.41.47.X	41.77.12.X
	-	41.221.84.X	41.59.7.X	41.87.10.X
2014	212.49.70.X	**196.0.17.X**	196.43.84.X	**41.221.97.X**
	212.49.95.X	196.0.26.X	196.45.144.X	41.87.10.X
	197.157.228.X	41.190.212.X	**41.77.228.X**	41.87.11.X
	197.211.12.X	196.0.64.X	196.41.43.X	41.77.14.X
	-	41.221.84.X	196.41.47.X	-
	-	41.84.196.X	41.59.7.X	-
2015	197.157.228.X	41.190.212.X	196.45.144.X	41.221.103.X
	212.49.70.X	196.0.64.X	**41.77.228.X**	41.77.14.X
	212.49.95.X	41.221.84.X	41.59.13.X	-
	196.202.202.X	212.88.112.X	196.41.47.X	-
	197.211.12.X	196.0.22.X	196.46.100.X	-
	-	41.190.197.X	196.45.159.X	-

NB: The IP in bold means they appeared more than twice in the studied years

Table 5.12: Unique SRC IP % Composition

Country	**# of packets**	**%**
Kenya	8,587	8.05
Tanzania	3,124	5.59
Uganda	900	1.99
Malawi	178	3.68
Total	12,789	19.31

of these countries rather than being concentrated as it is now.

5.12 Significance of Unique IP Addresses

As much as the countries have a different number of packets and unique IP addresses the rate at which each of these countries is exposed to IBR may not be proportional to it. Computing the percentage composition of unique IP addresses was done by calculating the total number of unique IP addresses that registered traffic divided by the total number of IP addresses of each country. The percentage calculation was not done by aggregating the total traffic in EAC rather per country. Table 5.12 shows a summary of the unique IP addresses registered and their percentage composition.

From Table 5.12 Kenya had the highest risk of IBR exposure than any of the three countries because the number of its unique IP addresses in relation to the total traffic it registered through all IP addresses, as a country, is high. This means that Kenya had the highest chance of malicious activity passing through the unallocated IP addresses. Despite not having the lowest packet count, Uganda had the lowest exposure rate to being exploited using IBR. That means that of the four countries under study, Uganda is more secure as far as unallocated addresses malicious activity is concerned.

It also means that the ration between unique IP addresses and the total number of IP addresses in Uganda is very small. The activities that were happening in Uganda's dark space reused the same IP address hence having a smaller percentage composition than the rest of the countries in EAC. The opposite of this happened in Kenya in that despite the fact that it registered more traffic, this traffic came from a wide range of unique IP addresses than any other country in EAC. Considering that IBR data contain packets that can provide insight of the network properties to the attacker, then given information present in Table 5.12, it is safe to assume that an attacker is more likely to learn more about Kenya than the rest of the countries because of it's vast number of packets.

5.13 Recommendations

IBR, like traffic in assigned IP addresses, uses port 445/TCP a lot more than any other port due to the nature of the services that it offers but also vulnerabilities attached to it. There is also a need for external firewalls to be set not to accept traffic on TCP ports like 135 or 445 for example. All services that are not in use ought to be disabled and all ports not in use closed, if not then the security updates need to be made more often with all applications patched. Most of the destination ports that were targeted provided access to sensitive services like file sharing, RPC or NetBIOS and carry sensitive unencrypted data.

By using a firewall an attacker would be prevented from sending messages to the workstation services. Most firewalls including Internet connection firewall block these ports by default. Both allocated and unallocated IP addresses need to be given the same amount of attention and protection as the damage caused by any of the two is equally bad. Lastly, if source IP addresses in IBR remain unattended or are not monitored they can repeatedly be used to perform malicious activities simply by changing the port of exploitation.

5.14 Summary

This chapter started by introducing what IBR is, how it works and identify the kind of device that can be used to identify network traffic in the dark-net. In Section 5.2, it went on further to explain and analyze the two variables that were identified as key elements that can be used to gain illegal access to networks. It identified ports that were more prevalent in various countries and attempted to explain the possibility of such ports being targeted because of the role they play. A table was made summarising the top four ports that were found in each of the four countries and how much each of them contributed to the total traffic.

The data expanded for a period of 72 months from 2009 to 2015 prompting the need to create top five ports that had the highest prevalence rate and were present in all the four countries during this period. Section 5.4 gave a detailed analysis of the source IP addresses and a list of top four IP addresses were listed together with its composition. The table also highlighted the IP addresses that were reused for at least once indicating continuity of malicious activity in these areas. A look at each of the four countries was made by the use of graphs and tables, with an explanation of how each of the ports

performed throughout the study period. This was a form of comparative analysis that focused on identifying patterns and trends.

In more ways than one IBR exhibited the same characteristics as traffic captured in assigned IP addresses as such it may be used in planning for security maintenance or even securing systems and computers from malicious activities that may, sooner than later appear in assigned IP addresses. This work has shown some promise in assessing the current and historical state of cybersecurity for countries based on the observed IBR emissions. Based on the observable traffic, one can apportion the likelihood of potential compromise or persistence of certain malware types within the IP address space of interest.

Chapter 6

Information Security Perspective: SHODAN

This chapter focuses on gaining security insight resulting from an analysis of data extracted from SHODAN. Each of the sections from Section 6.2 to 6.9 explains each of the variables that SHODAN was able to gather information on. A consideration of how OpenSSL can be used as a source of information that can exhibit vulnerability types and its associated CVE-IDs is explained in Section 6.2. The list is a long one as such only a sample of the OpenSSL versions that were found was shown so as to demonstrate the principle behind it. Such versions of OpenSSL were present in all the four countries.

In Section 6.3 the study shows network devices that used default passwords for their authentication in the East African Community (EAC) and were not properly protected. Section 6.4 shows and explains the operating systems that released their banner to SHODAN when a probe was made. Tables 6.3 and 6.4 show a summary of these findings. The products that were detected by SHODAN are shown in Section 6.5 which is followed by port prevalence in the EAC in Section 6.6. ISPs that offered Internet services to clients in the countries under study are shown in Section 6.7 together with their statistics. The last variable to be looked at is device type shown in Section 6.8. Looking back to Section 4.4 the study showed that only data for Malawi and Tanzania were available when extracting SHODAN data for 2015. Therefore Section 6.9 takes a look at these two countries to track the changes that have occurred within the 14 months period by looking at a subset of the variables that have been explained in this chapter. This section is followed by a summary which winds up the findings in Section 6.10.

6.1 Introduction

In Chapter 4 the study introduced SHODAN as a search engine on the Internet that grabs banners from the Internet of Things (IoT) connected devices around the world. It generates customised results in that in order to come up with specific outcome SHODAN uses filters to narrow down the content preferred by the user. Such results include but not limited to web cameras, phones, security systems, routers and all other devices in the IoT category that is not properly secured. It is also able to acquire geolocation of the devices it detects as was seen during data collection for this study.

In order to access banner of devices and system SHODAN uses ports 21/TCP(FTP), 22/TCP(SSH), 23/TCP(Telnet), and 80/TCP(HTTP) as the core ports for scanning of which port 80/TCP contains more traffic than the rest (Schearer, 2010). Worth noting and recalling back to Chapter 4 is the fact that the ports used in this Chapter are TCP ports. This chapter looked at a sample of these devices and other variables that are embedded within systems or devices. These variables are ports, operating systems, OpenSSL, products that were used by any of the devices detected, ISPs and configuration settings which may include security settings, poor passwords or default passwords on all devices and systems that are properly secured.

6.2 OpenSSL Versions Used in EAC

The SHODAN data contained a lot of variables including OpenSSL used in various products. As defined earlier in Section 2.5, OpenSSL is a cryptographic library used in many server products to secure communications against eavesdropping. SHODAN was able to pull banners from the products that used it for encrypting communication. Table 6.1 shows various versions of OpenSSL, the type of attack or vulnerability that the version is exposed to as well as the CVE-ID that depict this vulnerability in these versions of OpenSSL. This was data taken from data sets C, D, E and F. For more details on data sets refer to Table 4.7 in Chapter 4. The table is not a summary of the versions, neither is the list of vulnerability types an exhaustive one. The table only shows in principle how much data an attacker can acquire should the attacker gain access to SHODAN with the victim's details in it. By knowing the version of OpenSSL an attacker only needs to find the vulnerabilities that are associated with it and how to bypass it.

Table 6.1: OpenSSL Version vs CVE

OpenSSL version	Vulnerability Type	CVE ID
OpenSSL/0.9.8e	Bypass protection	CVE-2011-4109
OpenSSL/1.0.1e-fips	DoS Overflow	CVE-2016-2842
OpenSSL/0.9.7e-p1	DoS Overflow	CVE-2006-3738
OpenSSL/1.0.2f	DROWN attack	CVE-2016-0800
OpenSSL/1.0.1i	DoS attack	CVE-2014-3507
OpenSSL/1.0.0-fips	DoS attack	CVE-2016-6302
OpenSSL/0.9.8y	DoS Exec Code Overflow	CVE-2014-0195
OpenSSL/1.0.1p-freebsd	DoS attack	CVE-2014-3513
OpenSSL/FIPS	Bypass protection	CVE-2007-5502
OpenSSL/0.9.8e-fips-rhel5	CCS Injection	CVE-2014-0224

Some of the CVE-IDs include CVE-2011-4109[1] which allows remote attackers to have an unspecified impact by triggering the failure of a policy check. CVE-2016-6302[2] which allows remote attackers to cause DoS via a ticket that is too short. Then there was CVE-2014-0224[3] which does not properly restrict processing of *ChangeCipherSpec* messages allowing MiTM attackers to trigger use of a zero-length master key in certain OpenSSL to OpenSSL communications.

6.3 Default Password Usage

SHODAN data set I (see Table 4.7) was used to search for devices in EAC that used default passwords. *Query 1* in *Listing 4.1* was used to search for default passwords in the study region. Initially, the search was not specified to routers but when the general query was run it produced results of which about 98% were routers. For that reason, it was opted to drop the other devices like switches which only showed two occurrences in Tanzania and Uganda. All these devices had username *'cisco'* and password *'cisco'* designated for one-time use but the owner left it unchanged with level 15 privilege. Worth noting is the fact that these default passwords for routers had level 15 privilege i.e. administrative mode which enables an attacker to have full access to the router (Davis, 2008). Such privilege in a router is equivalent to having root privileges in Unix or administrator privileges in Windows. In Table 6.2 shows percentage distribution of default password used by the routers in EAC. SHODAN recorded 363 routers that contained default passwords. Kenya contributed more to this number with 61.3% followed by Uganda with 25.2%

[1]https://securityvulns.com/CVE-2011-4109.html
[2]https://access.redhat.com/security/cve/cve-2016-6302
[3]https://cve.mitre.org/cgi-bin/cvename.cgi?name=cve-2014-0224

Table 6.2: Router Default Password Distribution

Country	# of records	%
Kenya	255	61.3
Uganda	93	25.2
Tanzania	42	10.8
Malawi	11	2.7
Total	402	100

then Tanzania lastly Malawi with 10.8% and 2.7% respectively. SHODAN also recorded default passwords for switches and three other unknown devices, but the number was not as significant as that of routers.

6.4 Detected Operating Systems

Each of the devices has an operating system that makes it work and while that is the case, each of the operating systems has its own vulnerabilities and strengths. In this section, this study focused on identifying operating systems that were present in the data that was collected by SHODAN. Table 6.3 is a summarised version of all the operating systems that SHODAN was able to pick up in all the four countries with data taken from data sets C, D, E, and F.

The number of operating systems varied in each country and so did its composition. For Kenya, Tanzania, and Uganda it was Ubuntu that was more dominant than any other operating system. It is an open source operating system that runs on network servers. For Malawi, however, it was CentOS, a community-supported distribution derived from sources freely provided to the public by Red Hat for Red Hat Enterprise Linux (RHEL). In the case of Malawi, Ubuntu came second with 29% of the total operating systems that were detected. Scientific Linux[4], an enterprise Linux rebuild sponsored by Fermi National Accelerator Laboratory with a bias toward High Energy and High-Intensity Physics community was only found in Kenya alone. Other operating systems that were not present in all the four countries include ClearOS (not present in Malawi) and Linux/SUSE (not present in Tanzania).

[4]https://www.scientificlinux.org/

Table 6.3: Operating System Identified by SHODAN

	%			
Operating System	**Kenya**	**Tanzania**	**Uganda**	**Malawi**
Ubuntu	33.40	33.66	41.26	29.99
CentOS	27.44	27.67	17.34	38.67
Debian	11.72	9.53	9.79	9.64
Windows	7.14	8.16	12.37	6.75
Unix	2.81	8.47	2.77	1.45
FreeBSD	6.45	1.73	2.28	0.48
Red Hat	3.16	6.10	6.02	2.29
Fedora	2.47	4.03	1.57	7.71
ClearOS	0.08	0.44	0.07	-
Scientific Linux	0.18	-	-	-
Oracle Linux[5]	4.22	0.18	2.50	2.89
Linux/SUSE	0.92	-	4.04	0.24

6.5 Products Detected by Shodan

The name of the product that generated the banner for SHODAN was one of the variables of consideration in this research. One of the main reasons for this option was the fact that every product that SHODAN pulled data from had a list of known vulnerabilities associated with it. It works more efficiently to the attacker as an error message would give insight to the attacker thus posing a risk to the security of system and applications. Data for this section was taken from data sets C, D, E, and F.

Table 6.4 shows a sample list of the products that were captured by SHODAN from all the four countries and on the right hand side of the table shows a list of vulnerabilities that were identified by both the attacker and the vendor of the product but in rare cases may not be known to the owner of the company using the products. There were over 300 unique products whose details were captured by SHODAN, all of which had multiple instances in each country under study. For examples, Microsoft Internet Information Services (MIIS), there were six version of it from Microsoft-IIS/6.0 to Microsoft-IIS/8.5 each with a list of known vulnerabilities. Table 6.4 shows CVE-ID for Microsoft-IIS/6.0 which causes a buffer overflow in Microsoft-IIS 5.0, 5.1, and 6.0 allows local and possibly remote attackers to execute arbitrary code via crafted Active Server Pages (ASP). This is one of the many vulnerabilities associated with this version of IIS; in version 8.5 there are vulnerabilities too.

Table 6.4 also shows other products from various vendors including FreeBSD, Apache, HP, Linksys VMware, Cisco, and Apple. It is a long list of products and vendors but the

principle behind this was to demonstrate what information can be acquired from these products.

Table 6.4: Product vs Vulnerability

Product	Vulnerability ID
Microsoft-IIS/6.0	CVE-2006-0026
Microsoft Exchange 2010 smtpd	CVE-2010-1690
Microsoft SQL Server	CVE-2012-1856
FreeBSD	CVE-2014-3954
Apache Tomcat/Coyote JSP engine	CVE-2015-5351
Linksys WRT54GL wireless-G router http config	CVE-2008-1247
Cisco catalyst switch telnetd	CVE-2001-0041
Cisco ASA 5510 firewall http config	CVE-2014-8730
Apple AirPort	CVE-2010-1804
VMware Authentication Daemon	CVE-2009-3707

6.6 Port Prevalence

These are some of the ports that were found open by SHODAN in data sets C, D, E and F. Once ports like 80/TCP (HTTP), 21/TCP (FTP), 22/TCP (SSH), 23/TCP (TELNET), 161/UDP (SNMP) and 5060/UDP (SIP) are left open, it enables SHODAN's capability to pull service banners from such devices and servers from the web. Some ports appear on both TCP and UDP like 443 for example.

Table 6.5 shows port prevalence in EAC countries with most of the ports that SHODAN uses to pull banners from. Some of the ports appear there because they were left open and have the potential to cause security breach while ports like port 21/TCP, 80/TCP, 443/TCP and few others are the primary ports that SHODAN uses should they be found open and unguarded to acquire data from all devices. A consideration is first made to those ports that SHODAN uses as core ports for banner extraction. With reference to the opening statement of this section, but also in Chapter 5, it can be seen that of those ports were available and were used by SHODAN in all of the countries under study. i.e port 80/TCP, 21/TCP, 22/TCP, 23/TCP, 161/UDP, and 5060/UDP. Port 5060 was not shown in the picture because its prevalence was not as dominant as the rest of the ports appearing on the table. Due to the nature of services run by other ports they have to be opened like port 80/TCP for example while for others they have to be traded carefully like port 135/TCP, 137/TCP, port 445/UDP, 111/TCP and 3389/TCP all of which were picked by SHODAN in data sets A, B, C, D, E and F.

Table 6.5: Top 10 Port Prevalence by Country

	% ports by Country			
Port Number	**Kenya**	**Tanzania**	**Uganda**	**Malawi**
Port 443	10.40	12.14	7.97	21.61
Port 80	16.46	30.18	13.51	16.50
Port 23	13.33	20.20	7.77	13.48
Port 22	8.92	18.85	11.28	14.55
Port 161	3.88	5.03	18.43	3.31
Port 21	5.49	9.30	1.28	10.02
Port 500	3.97	6.76	3.04	4.47
Port 2000	5.27	20.42	2.70	20.25
Port 4500	2.70	4.55	1.72	4.27
Port 8080	4.85	7.19	1.95	1.51

Total number = 72132

Of these core ports, port 80/TCP was used more than any other port followed by port 23/TCP and then port 22/TCP which was followed by port 21/TCP and lastly port 161/UDP. In Kenya and Tanzania port 80/TCP was more dominant than any other port, in Uganda the most dominant one was port 161/UDP while in Malawi it was port 443/TCP and 443/UDP. Based on the protocol that each of these ports uses an attacker may explore more along those lines. More details of the raw data are shown in Table 5 Appendix B. It was not possible to rank all the ports in all the four countries using the same measure because the ports that were more dominant in one country are not the same for another. For examples from Table 6.5, Port 443/TCP is dominant in Malawi (21.61%) while this is not the case for the other three countries. For Kenya and Tanzania the dominant one is Port 80/TCP (16.46%) and for Uganda, it is Port 161/UDP.

6.7 ISP Distribution

The ISP provides organizations with the IP space for any device that accesses the Internet making a variable of interest in this study. Table 6.6 show a list of top ISPs in EAC and how many of their clients IP addresses showed up on SHODAN's raider. All the top ten ISPs contributed to at least 80% of the total IP addresses found by SHODAN. In Kenya, One Communications Ltd had the highest number of IP addresses registered under it (16.76%). The coordinates of One Communications Ltd[6] showed that it is located in 13651-00800 Westlands, Nairobi, Kenya followed by Wananchi group, a leading business

[6]https://www.safaricom.co.ke/

providing affordable entertainment and connectivity for the rapidly growing middle class in East Africa.

In Tanzania, Tanzania Telecommunications Company Limited[7] (TTCL) the oldest and largest ISP in Tanzania had the highest number of IP addresses. It also showed the highest prevalence of its clients on SHODAN whose IP addresses were more than twice as much as its closest run which was Startel (T) Limited as it can be seen from Table 6.6. The same pattern goes for Uganda where MTN Uganda[8] had at least twice the number of IP address and the second most dominant ISP, with MTN Uganda having 30.22% of the total number of ISPs and Roke Investments International Ltd had 15.01%. The biggest margin is seen in Malawi with Skyband Corporation Ltd[9] contributing 44.49% of the total IP addresses registered by SHODAN and Malawi Telecommunications Ltd (MTL) coming second with less than half of what Skyband had. Note that each of these countries had over ten ISPs but only the top five of them have been shown here. The rest of the data sets for ISP have been appended to Appendix B. All this was picked by SHODAN in data sets C, D, E, and F.

6.8 Detected Devices

Data for this section was taken from data sets C, D, E and F with Table 6.7 showing devices that were detected by SHODAN, meaning that it (SHODAN) grabbed their banners because they were not secure. A properly secured device is not supposed to be detected by SHODAN. This table shows results covering all the four countries under study. For proper representation percentage was used. Worth noting is how these devices appeared in each of the countries. All of the devices appeared in all the countries except for media devices, web cameras, and printers. The dash in the table means that no device of that type was present in that country.

Apart from WAP in Malawi, a router was the device that was dominant throughout the four countries. WAP is technical standard for transferring documents, especially in web pages, over a computer network to cellular phones and other handheld wireless devices i.e it is a protocol for accessing information over a mobile wireless network. Only Kenya had the lowest level of WAP devices detected. With Africans migrating to mobile banking this

[7]https://www.ttcl.co.tz/
[8]https://www.mtn.co.ug/en/Pages/default.aspx
[9]https://www.skyband.mw/

Table 6.6: Top 5 ISPs by County and their % Distribution

Country	Rank	ISP	%
Kenya	1	One Communications Ltd	16.76
	2	Wananchi Group	14.31
	3	Access Kenya Group Ltd	13.87
	4	Jamii Telecommunications Limited	11.72
	5	Communication Solutions Ltd	8.49
Sum			**65.15**
Tanzania	1	Tanzania Telecommunications Co. Ltd	19.65
	2	Startel (T) Ltd	9.45
	3	Simbanet (T) Ltd	6.98
	4	Habari Node Ltd	6.63
	5	Spice Net Tanzania Ltd	5.91
Sum			**48.62**
Uganda	1	MTN Uganda	30.22
	2	Roke Investments International Ltd	15.01
	3	Uganda Telecom	10.11
	4	Africell Uganda Ltd	7.30
	5	Research and Education Network of Uganda	4.28
Sum			**66.92**
Malawi	1	Skyband Corporation Ltd	44.49
	2	Malawi Telecommunications Ltd (MTL)	16.59
	3	globe Internet limited	10.95
	4	Roya Hosting LLC	5.27
	5	Airtel Malawi network	5.14
Sum			**82.44**

may be a big threat to that development as it provides an easy means to which attackers can gain information. It was the second most detected device by SHODAN after routers

Another device worth noting was the presence of web cameras in the data set. The numbers were small but the fact that they appeared shows how vulnerable they are once they are connected to the Internet thereby providing the attacker with illegal means of surveillance on the owner of the web camera. Security misconfigurations were also detected in all four countries with Uganda recording the highest number of them. Security misconfiguration is among the top security risks of web applications. Application misconfiguration attacks exploit configuration weaknesses found in web applications. Misconfiguration is defined as configuration mistakes that result in unintended application behavior that includes misuse of default passwords, privileges, and excessive debugging information disclosure.

After routers and WAP, another common device that was detected was a firewall. Often

Table 6.7: Detected Devices in EAC

	% composition by Country			
Device-type	**Kenya**	**Tanzania**	**Uganda**	**Malawi**
Switch	1.64	1.22	9.96	0.21
WAP	18.25	9.61	18.43	57.39
Firewall	9.49	11.40	19.07	2.55
PBX	0.42	1.14	1.69	0.21
Router	68.06	75.57	46.82	37.73
Media device	0.03	0.16	-	-
web-cam	1.19	0.24	-	0.32
Printer	0.03	0.24	-	0.32
security-misc	0.87	0.41	4.03	1.28

defined as a network security system, either hardware or software-based, that controls incoming and outgoing network traffic based on a set of rules, firewalls were present in all four countries with Tanzania having more instances. There was no specific definition of what media device this was but it appeared in Kenya and Tanzania. Another device least expected to be found was private branch exchange (PBX). A PBX is a telephone system within an enterprise that switches calls between users in the enterprise on local lines while allowing all users to share a certain number of external phone lines. It was present in all the countries and was more dominant in Uganda and least in Malawi. There were also instances where printers were detected but their cases were very minimal and there were no instances of this in Uganda.

6.9 Tracking Changes in SHODAN Data

The study managed to have access to SHODAN data for June 2015 with data for this section taken from data sets A, B, D, and E. Unfortunately, the data did not accommodate countries like Kenya and Uganda thus this section only covers differences that were observed in Malawi and Tanzania. The period was chosen because it was long enough to track any changes in the variables. In this section three variables will be looked into, namely: ports, ISP, and device types detected. For clarity purposes, visual aid in the form of graphs was used especially for Port numbers and ISP.

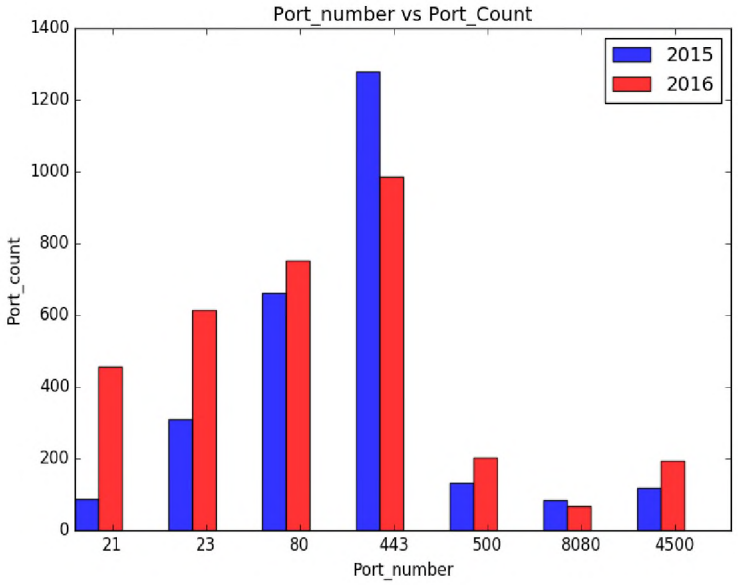


Figure 6.1: Port Prevalence in Malawi (2015-2016)

6.9.1 Port

SHODAN picked up a lot of ports with multiple instances from both Tanzania and Malawi data sets. This study only selected ports that showed more instances than others. All of these ports SHODAN found them either open or not properly guarded otherwise they would be appearing in the data sets. More of the content of the data sets check-in Appendix B which contain all the data that was used for this research.

Figure 6.1 shows port prevalence in Malawi from June 2015 to August in 2016. The patterns and trends of ports that SHODAN used to extract data were the same and so was that of ports that seemed to pose potential weakness to their devices. In 2015, port 443/TCP and 443/UDP showed the highest port count than any other port followed by port 80/TCP. Then port 23/TCP was followed by port 21/TCP then port 500/TCP, port 500/UDP), port 8080/TCP and lastly port 4500/udp. Figure 6.1 also shows that only port count for port 443/TCP and port 8080/TCP declined while the rest of the ports increased their count, with port 21/TCP showing a big margin between what was recorded by SHODAN in 2015 and that of 2016 followed by port 23/TCP.

Figure 6.2 shows port prevalence in Tanzania covering a thirteen months period (June 2015 - August 2016). Unlike Malawi, Tanzania had port 80/TCP used for HTTP and it was the dominant port that SHODAN picked up as more vulnerable than any other port. This is followed by port 23/TCP which is used for SMTP, port 443/TCP and 443/TCP

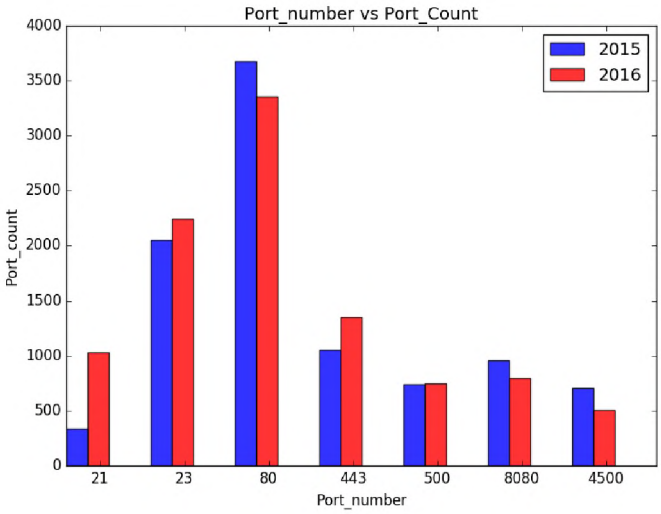


Figure 6.2: Port Prevalence in Tanzania (2015-2016)

used for HTTP over SSL comes third. This pattern is consistent for both 2015 and 2016. After port 443 the pattern changed to port 21/TCP which was used for FTP being fourth in 2016 but was last in 2015 among the top ports and then port 8080/TCP which works the same way as port 80/TCP followed.

6.9.2 Internet Service Providers

Figure 6.3 and 6.4 show ISPs in both Malawi and Tanzania respectively for both 2015 and 2016. The ones shown in these graphs are the top ISPs found in these countries for the study period. Worth noting is that there are more ISPs in the data set with some only appearing in either 2015 or 2016. For example, in 2016 two more ISP were detected i.e Simbanet-malawi and Royal Hosting LLC.

From Figure 6.3, all of the five ISP increased the number of clients in their network based on the increased number of instances. Each of these has an IP address attached to them and in some cases together with the MAC address and their coordinates. Like in 2015, Skyband dominated vulnerable clients in 2016 followed by MTL then Globe. There was a twist for Airtel and Broadband wherein 2015 there were fewer Airtel clients compared to Broadband but the reverse is true for 2016.

Such changes were not the same for Tanzania. Figure 6.4 showed that in 2015, TTCL had more of its clients detected by SHODAN in 2016 than any other ISP. From there onward, the pattern changes completely, with Startel being second in 2016 when it was fourth in

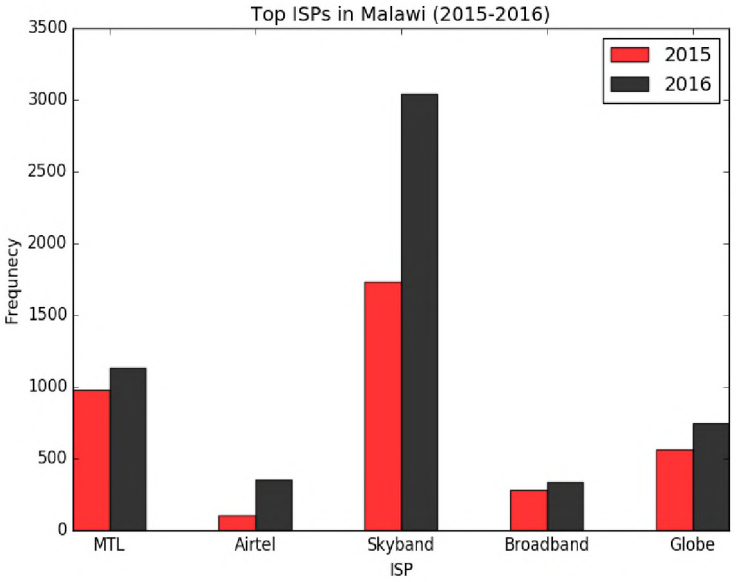


Figure 6.3: ISPs in Malawi

2015. Simbanet was pushed a step back to third from second while Hariba Node was last in the top ISPs but came fourth in 2016. SpiceNet which was third in 2015 came last in 2016.

6.9.3 Device Type

Table 6.8 shows names of devices that were detected by SHODAN between June 2015 to August 2016 in Tanzania and Malawi and the % change that occurred within the 14 months period. For those with a negative value, it means that there was a decrease in value, for those with positive value means there was an increase in the value while those with N/A it means it was not possible to compute their value because the initial value was not available. There were significant changes in all the devices except for media device in Tanzania where the number remained the same. Comparing Tanzania's 2015 data set to that of 2016 it showed that two more devices were introduced in the form of web-cam and security misconfiguration. In 2015 there were no web cameras in the data set but in 2016 three web cameras were detected. In addition to this, five security misconfigurations in systems were detected. There was a significant change in the number of switches detected for 2015 and 2016. In 2015, 414 switches were detected by SHODAN as compared to 15 in 2016. The same changes happened for firewalls (62 in 2015 to 140 in 2016), WAP (from 5 in 2015 to 118 in 2016), router (from 9 in 2015 to 928 n 2016). On the other hand, the number of PBX decreased from 35 in 2015 to 14 in 2016 and so did printers (from 117 in

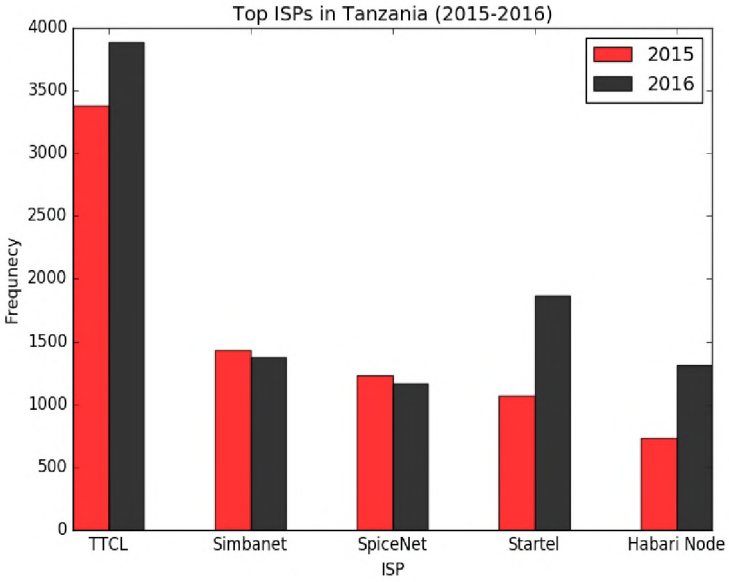


Figure 6.4: ISPs in Tanzania

2015 to 3 in 2016). Overall, there were more devices detected in 2016 than in 2015, this is in both the actual numbers but also the type of device detected.

For Malawi, the changes are equally big with only media device being the only missing device type in 2016. Like the case of Tanzania, the number of switches declined from 45 in 2015 to two in 2016. WAP devices were not detected in 2015 but in 2016, 540 of them were detected. The same goes for PBX, web-cams and security misconfigurations not present in 2015 but were available in 2016. In addition to this, 2016 dataset had 12 security misconfigurations detected in 2016, three web cameras and two PBX. The number of printers reduced from 23 to three. While there was only one router detected in 2015, 2016 dataset recorded 355 routers and there were 14 more firewalls that were detected in 2016 than those of 2015.

6.10 Summary

This chapter showed that SHODAN can pull banners from operating systems, devices, ports, products and even OpenSSL. From there it looked at a different operating system that SHODAN can pull data from, of which a table was made with over ten operating systems being detected. It also showed how certain devices whose default passwords are not changed can be detected once they are connected to the Internet rendering them vulnerable to attacks. Proper measures need to be taken to ensure that ports 443 (HTTPS), 80

Table 6.8: Device Types Used in Tanzania & Malawi

	Tanzania			Malawi		
Device- Name	**2015**	**2016**	**% change**	**2015**	**2016**	**% change**
Switch	414	15	-3.62	45	2	-4.44
WAP	5	118	2360.00	-	540	N/A
Firewall	62	140	225.81	10	24	240.00
PBX	35	14	-40.00	-	2	N/A
Router	9	928	10311.11	1	355	35500.00
Media device	2	2	0	1	-	-100
web-cam	-	3	N/A	-	3	N/A
Printer	117	3	-2.56	23	3	-13.04
security-misc	-	5	N/A	-	12	N/A
	644	1228	190.68	80	941	1176.25

(HTTP), 21 (FTP), 22 (SSH), 23 (TELNET), 161 (SNMP) and 5060 (SIP) are properly secured for fear of being used as a means of further access into the networks. These port among others have proved to be key entry point from which SHODAN can pull banners from.

Considering that there was a 14 months gap between the two data collection points, an entire section was used to elaborate the changes that happened between the two periods. The data present was only that of Malawi and Tanzania which formed data sets A, B, D and E, as such all explanation made in Section 6.9 focused on these two countries. Graphs were plotted using python to show these changes. In cases where graphs were not appropriate, tables were used to show the changes. Device type, ports, and operating systems were the three variables that were looked at when tracking these changes. Data from which this chapter is based was added to *Appendix B*. An extension of SHODAN data analysis is shown in Chapter 7 building on this work.

Chapter 7

Quantifying Vulnerabilities in EAC Using Shodan Data

This chapter extends on the vulnerabilities found in Chapter 6 that enable quantification of variables by picking instances that appeared in the SHODAN data set G and H shown in Section 4.4, Table 4.7. These datasets were picked because they contained more instances than the other vulnerabilities found in EAC. It explains the principle behind the CVSS scoring system and how it can be used to detect vulnerabilities and each can be quantified. It shows how these vulnerabilities are computed.

Among other issues addressed in this chapter include a brief explanation of Common Vulnerability Exposures (CVE) in Section 7.2 and how it used in line with CVSS. It goes on further to explain the components that constitute CVSS formula in Section 7.3. Further explanation on how to compute components on the score are described in Section 7.4. Once the computation is done, the chapter goes further to reflect on the significance of the computation made for CVE- 2015-0204 and CVE-2014-0160 to each of the countries in EAC in Section 7.5. Considering that two of the sample vulnerabilities used OpenSSL, Section 7.6 explains and shows how OpenSSL was usage in EAC during the study period.

7.1 Introduction

One of the major problems why Africa as a continent is finding hard to curb the issue of cybersecurity vulnerabilities and threats is because it has not been able to quantify certain aspects of security. For instance, in East Africa, no record of quantification of the

vulnerabilities that systems are vulnerable to has been documented and published. Thus there is no progress towards mitigating such vulnerabilities and threats. Using SHODAN data set G and H, this chapter addressed some of the principles that show the possibility to attain this goal i.e. quantifying security vulnerabilities.

One measurement that has been extensively used before in vulnerability assessment is CVE metric (Swart *et al.*, 2014). CVE provides a structured means through which information security vulnerabilities can be exchanged and makes data sharing across separate platforms easier due to its common naming (Mitre, 2015). This technique helps in scoring and assessing the seriousness of the vulnerabilities available on a network, more importantly by looking into the severity of the vulnerabilities.

Information about the type of vulnerability, time stamp, and severity score is provided by the CVE entries (Bozorgi, Saul, Savage, and Voelker, 2010). The severity and risk of the identified vulnerabilities are quantified using CVSS which is an open framework (Houmb, Franqueira, and Engum, 2010). Base metrics, a component of the three metrics of CVSS, captures fundamental vulnerability features such as access vendor, access complexity, authenticity, and the impacts on confidentiality, integrity, and availability (Fruhwirth and Mannisto, 2009).

7.2 Common Vulnerability Exposures Identity

To compute the score for any vulnerability there is need to have Common Vulnerability Exposures Identity (CVE-ID) which is a primary key for such a computation. This study focused on a sample of the CVE-IDs that were detected by Shodan in EAC i.e. CVE-2014-0160 often referred to as Heartbleed and CVE-2015-0204 also know as FREAK Vulnerability. The analysis of this chapter is based on data set G and H shown in Table 4.7 as a summary.

The Freak vulnerability allows attackers to grab HTTPS connections between vulnerable clients and servers and force them to use weakened encryption which can then be decrypted or altered (VanderSloot, Amann, Bernhard, Durumeric, Bailey, and Halderman, 2016). A malicious server could make a TLS/SSL client prone to this, using OpenSSL, use a weaker key exchange method i.e. it allows an attacker to intercept HTTPS connections between vulnerable clients and servers and facilitate brute-force decryption (weakened encryption) (RedHat, 2015b). This weak encryption makes it easier for an attacker to perform a brute force. CVE-2015-0204 affects all operating system platforms. For example, Linux systems

like Ubuntu, Debian, CentOS, RedHat and Fedora (RedHat, 2015a) and all computers and servers running on the Windows family of operating systems both client and server side are also affected by FREAK Vulnerability (EventTracker, 2015).

On the other hand, the Heartbleed vulnerability allows an attacker on the Internet to read the systems protected memory by using vulnerable versions of the OpenSSL software (via crafted packets that trigger a buffer over-read) (Rainie, Duggan, and Tyson, 2014). This allows stealing of information protected by the TLS encryption used to secure the systems on the Internet. This occurs due to the fact that more data can be read than it should be allowed to, a situation referred to as buffer over-read. This compromises the secret keys used to identify the service providers and to encrypt the traffic which contains the names and passwords of the users involved and the actual content of the communication (Ghafoor, Jattala, Durrani, and Tahir, 2014). This compromise allows attackers to create fake certificates and eavesdrop on communications, using a vulnerable OpenSSL instance for TLS as a server or a client (it leads to the leakage of memory contents from the client to the server and vice versa) and steal data directly from the services and users and to impersonate services and users (Lee, Yi, Tan, Goh, Lee, and Yeo, 2014).

What makes Heartbleed more lethal is that its exploit is possible even when using a user without administrative privileges and that its attacks occur without leaving a trace (Heartbleed, 2014). Some operating system distributions such as Debian Wheezy, Ubuntu 12.04.4 LTS, CentOS 6.5, Fedora 18, among others, were shipped with the potentially vulnerable OpenSSL version (1.0.1) (Durumeric *et al.*, 2014). A patch was developed for this vulnerability but website owners are slow to secure their systems through patching them and updating their certificates. Proof of concept exploit code is available online in circulating often being used by cybercriminals (Ghafoor *et al.*, 2014).

CVE-ID provides key needed to compute CVSS score as the CVE databases attaches to it all the characteristics that are associated with it acquired from the CVE dictionary. Thus CVE-ID is the primary key to all data needed.

7.3 Common Vulnerability Scoring System

In Section 2.9 we referred to CVSS as an open framework for scoring the risk associated with vulnerabilities found in either software, hardware or firmware vulnerabilities (Ali, Zavarsky, Lindskog, and Ruhl, 2011a). In this section, we take a deeper route in using the CVSS framework in order to quantify vulnerabilities. These vulnerabilities can pose

a critical risk to any organization operating a computer network and can be difficult to categorize and mitigate. It provides a way to capture the main attributes of a vulnerability, and produce a numerical score reflecting its severity and textual representation of that score. The numerical score can then be translated into a qualitative representation to help organizations properly assess and prioritize their vulnerability management processes. Its wide usage has been acknowledged by well-established vendors like CERT, Cisco, Union Pacific, Microsoft and Symantec and it is preferred as it provides accurate and consistent vulnerability impact scores (Mell, Kent, and Romanosky, 2007).

7.3.1 CVSS Break Down

The CVSS uses an algorithm to determine three severity rating scores: Base, Temporal and Environmental. The scores are numeric; the sum of these categories range from 0.0 to 10.0 with 10.0 being the most severe (Ali *et al.*, 2011a). A CVSS score is made up of three possible metric groups shown in Figure 7.1. The three groups as shown in (Ali *et al.*, 2011a) are:

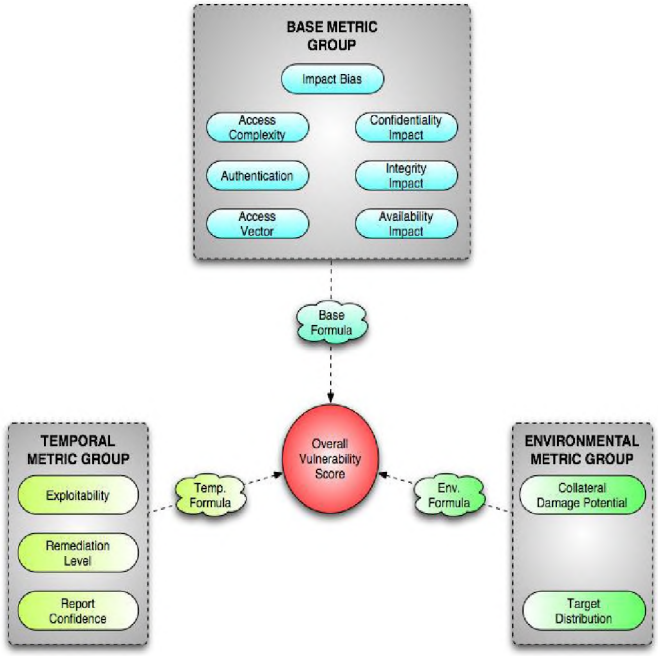
1. Base metric group: It is a mandatory Score by vendor or analyst and the metric most relied upon by enterprises and deals with the inherent qualities of a vulnerability.
2. Temporal metric group: Optional score by vendor or analyst representing the qualities of the vulnerability that change over time
3. Environmental metric group: Optional score by end-user representing the qualities of the vulnerability that are specific to the affected user's environment

The FREAK vulnerability in this study had base metric values of Access Vector [AV]: Network [N], Access Complexity [AC]: Medium [M], Authentication [Au]: None [N], Confidentiality Impact [C]: Partial [P], Integrity Impact [I]: Partial [P], Availability Impact [A]: Complete [C]. It also had temporal metric values of Exploitability [E]: Functional [F], Remediation Level [RL]: Official-fix [OF], Report Confidence [RC]: Confirmed [C]. Lastly the environmental metric values Collateral Damage Potential: None [N] - High [H], Target Distribution [TD]: None [N] - High [H], Confidentiality Req [CR]: Medium [M], Integrity Req [IR]: Medium [M], Availability Req [AR]: Medium [M]. Table 7.1 is a summary of base, temporal and environmental vectors for CVE-2015-0204.

The principle behind CVSS is to enable organizations to have the ability to prioritize which vulnerabilities to fix first and gauge the impact of the vulnerabilities on their

Table 7.1: Base, Temporal and Environmental Vectors for CVE-2015-0204

Metric group	Vector
Base	AV:[N]/AC:[M]/Au:[N]/C:[P]/I:[P]/A:[C]
Temporal	E:[F]/RL:[OF]/RC:[C]
Environmental	CDP:[N,H]/TD:[N,H]/ CR:[M]/ IR:[M]/AR:[M]



json

Figure 7.1: CVSS metric groups (Schiffman, 2005)

systems based on the environment in which they are in. Two different organizations can be affected by the same vulnerability but react to it differently depending on how such a vulnerability affects their industry.

7.4 Computation of CVSS Score Using CVE-ID

In this section the study used data set G and H to look the scores for CVE-2015-0204 and CVE-2014-0160 can be computed and then verify it with the countries under study to see the extent to which they are affected by these two vulnerabilities. The SHODAN data collected and analyzed showed that CVE-2015-0204 and CVE-2014-0160 are among the vulnerabilities that were found in EAC. The computation was done in phases starting with base metric group, then temporal metric group followed by the environmental metric group.

Table 7.2: CVE-2015-0204 Base Metric Values

Base Metric	Evaluation	Score
Access Vector (AV)	Network(N)	1.00
Access Complexity(AC)	Medium(M)	0.61
Authentication (Au)	None(N)	0.704
Confidentiality Impact(C)	Partial(P)	0.275
Integrity Impact (I)	Partial(P)	0.275
Availability Impact(A)	Complete(C)	0.660

7.4.1 Computation of Base Score

Using data set H shown in Tables 4.6 and 4.7 respectively, the study examined CVE-2015-0204 using the model used in (Houmb *et al.*, 2010). Since the vulnerability can be exploited remotely, the Access Vector was 'Network'. The Access Complexity was 'Medium' because it allows an attacker to intercept HTTPS connections between vulnerable clients and servers in order for this exploit to be successful; the attacker only needs to access a vulnerable client since the scope of this CVE is only client code based. A successful exploit of this vulnerability can lead the attacker to break or steal or manipulate sensitive data. No authentication is required to trigger the vulnerability since any Internet user can connect to the web server, so the Authentication metric was 'None'.

The description of CVE-2015-0204 showed that the vulnerability is exploitable by executing arbitrary code which intercepts HTTPS connections between vulnerable clients and servers and facilitates brute-force decryption, thereby altering web content and possibly viewing local user or configuration information like connection settings and passwords to back-end databases, thus the confidentiality and integrity Impact metrics for this vulnerability were set to 'Partial'. These were not set to 'Complete' because it is possible that the attacker has no intention of altering the content or leaking it hence setting to a minimum potential.

It was also possible to assume the worst case scenario, using the same vulnerability where the attacker completely disrupts the integrity and confidentiality of the content in which case the study set the evaluation to 'Complete' and the score will be entirely different. This is to say that depending on what the attacker is up to, the score may change from time to time and the algorithm for this computation provides room for such dynamic scenarios. For demonstration purposes the worst case scenario was not assumed, rather it was evaluated to 'Partial'. A summary of all base metric values of CVE-2015-0204 are shown in Table 7.2.

Listing 7.1: Formula for computing Base Score

```
BaseScore = round_to_1_decimal(((0.6*Impact)
          +(0.4*Exploitability)1.5)*f(Impact))

Impact = 10.41*(1-(1-ConfImpact)*(1-IntegImpact)
       *(1-AvailImpact))

Exploitability = 20 * AccessComplexity * Authentication
               * AccessVector

f(Impact) = 0 if Impact=0; 1.176 otherwise

AccessComplexity = case AccessComplexity of
                        high: 0.35
                      medium: 0.61
                         low: 0.71

Authentication = case Authentication of
                 No authentication:                        0.704
                 Single instance of authentication:        0.56
                 Multiple instances of authentication: 0.45

AccessVector     = case AccessVector of
                        Requires local access:    0.395
                        Local Network accessible: 0.646
                        Network accessible:       1

ConfImpact       = case ConfidentialityImpact of
                        none:             0
                        partial:          0.275
                        complete:         0.660

IntegImpact      = case IntegrityImpact of
                        none:             0
                        partial:          0.275
                        complete:         0.660

AvailImpact      = case AvailabilityImpact of
                        none:             0
                        partial:          0.275
                        complete:         0.660
```

Listing 7.2: Computation of FREAK attack base score

```
1  Impact = 10.41 * (1-(1-ConfImpact) * (1-IntegImpact)
2           * (1-AvailImpact))
3         = 10.41*(1-(1-0.275)*(1-0.275)*(1-0.660))
4         = 10.41*(1-(0.725)*(0.725)*(0.34))
5         = 10.41*(1 - 0.179)
6         = 10.41*0.82
7         = 8.54
8
9  Since Impact is not equal to zero
10 Therefore f(impact) = 1.176
11
12 Exploitability = 20 * AccessComplexity
13                  * Authentication * AccessVector
14                = 20*0.61*0.704*1
15                = 8.59
16
17 BaseScore = round_to_1_decimal(((0.6 * Impact)
18              +(0.4*Exploitability)1.5)*f(Impact))
19 BaseScore =((0.6*8.54) + (0.4*8.59) - 1.5))*1.176
20           =(5.124 + 1.866)*1.176
21           = 7.06*1.176
22           = 8.3
```

If the vulnerability was exploited to cause DoS, the Availability Impact was set to 'Complete'. Since this is the highest possible base score of the exploitation options, it is used as the base score.

The base vector for this vulnerability was, therefore: AV:N/AC:M/Au:N/C:P/I:P/A:C. Now that all the necessary base metrics were identified the base score of FREAK attack was computed as follows:

The study also examined CVE-2014-0160 using dataset G: Heartbleed vulnerability is exploited remotely making its Access Vector to be 'Network'. The Access Complexity was set to 'Medium' because no additional circumstances need to exist for this exploit to be successful; the attacker only needs to craft a proper exploit script for an exploit to be successful. No authentication was required to trigger the vulnerability since any Internet user can connect to the web server, so the Authentication metric was 'None'. A successful exploit of this vulnerability allows stealing the information protected, under normal conditions, by the SSL/TLS encryption used to secure systems on the Internet.

Since the vulnerability can be exploited using multiple methods with different outcomes,

Table 7.3: CVE-2014-0160 Base Metric Values

Base Metric	Evaluation	Score
Access Vector (AV)	Network(N)	1.00
Access Complexity(AC)	Low(L)	0.71
Authentication (Au)	None(N)	0.704
Confidentiality Impact(C)	Complete(C)	0.660
Integrity Impact (I)	Complete(C)	0.660
Availability Impact(A)	Complete(C)	0.660

a method that could provide the highest score in which case a vulnerability is exploited to cause a DoS was picked, the Availability Impact was set to 'Complete'. If the vulnerability was exploited to execute arbitrary code thereby altering the names and passwords of the users and the actual content or configuration information like connection settings and attack targets memory buffers then each of the Impact metrics had to be set to 'Complete' because of the possibility of a complete system compromise. A summary of all base metric values of CVE-2015-0204 is shown in Table 7.3. The base vector for Heartbleed vulnerability was, therefore: AV:N/AC:L/Au:N/C:C/I:C/A:C. The score for CVE-2014-0160 was computed as shown in Listings 7.2 and 7.3:

7.4.2 Computation of Temporal Score

Depending on the vulnerability under study, the temporal equation may or may not be applied to the base score so that the score is properly adjusted to reflect the changes made to the vulnerability. It is very important to note that temporal equation only applies to a variable that changes over time. If employed, the temporal equation combines the temporal metrics with the base score to produce a temporal score ranging from 0.0 to 10.0. In addition to this, it is also worth noting that the temporal score produces a score which is not less than 33% lower in relation to the base score from which it is based and no higher than its base score. The temporal score can be summarised as a way in which vendors provide feedback to their clients since it offers metrics that aim at getting feedback but also mitigate the vulnerabilities reported. The temporal equation was given as shown in Listing 7.4:

With CVE-2015-0204 under review, the research looked at the temporal metric values that apply to it. CVE-2015-0204 exploit code is known to exist and therefore exploitability metric value was set to 'Functional' since it is in existence and operational. Microsoft released patch MS15-031 for this vulnerability and so the Remediation Level was 'Official-Fix'. For Linux systems, updates to mitigate FREAK Vulnerability are also available and

Listing 7.3: Computation of Heartbleed attack base score

```
Impact = 10.41*(1-(1-ConfImpact)*(1-IntegImpact)
          *(1-AvailImpact))
       = 10.41*(1-(1-0.660)*(1-0.660)*(1-0.660))
       = 10.41*(1-(0.34)*(0.34)*(0.34))
       = 10.41*(1-0.039)
       = 10.41*0.96
       = 9.99

Since Impact is not equal to zero
Therefore f(impact) = 1.176

Exploitability = 20 * AccessComplexity * Authentication
                 * AccessVector
               = 20 * 0.71 * 0.704 * 1
               = 10.0

BaseScore = round_to_1_decimal(((0.6*Impact)
            +(0.4*Exploitability)1.5)*f(Impact))
BaseScore =((0.6*9.99) + (0.4*10.0) - 1.5))*1.176
          =(5.994 + 2.5)*1.176
          = 8.494
          = 9.989
To one decimal place = 10
```

Listing 7.4: Formula for computing Temporal Score

```
1  TemporalScore = BaseScore * Exploitability * RemediationLevel
2                  * ReportConfidence
3  Exploitability  =  case Exploitability of
4                          unproven:              0.85
5                          proof-of-concept:      0.9
6                          functional:            0.95
7                          high:                  1.00
8                          not defined            1.00
9
10 RemediationLevel = case RemediationLevel of
11                         official-fix:          0.87
12                         temporary-fix:         0.90
13                         workaround:            0.95
14                         unavailable:           1.00
15                         not defined            1.00
16 ReportConfidence = case ReportConfidence of
17                         unconfirmed:           0.90
18                         uncorroborated:        0.95
19                         confirmed:             1.00
20                         not defined            1.00
```

Table 7.4: CVE-2015-0204 Temporal Metric Values

Temporal Metric	**Evaluation**	**Score**
Exploitability (E)	Functional(F)	0.95
Remediation Level(RL)	Official-fix(OF)	0.87
Report Confidence(RC)	Confirmed (C)	1.00

often done by either upgrading OpenSSL or disable EXPORT -grade ciphers in your client or server. The patch released also ensured that report confidence was set to 'Confirmed'. The temporal vector for this vulnerability is, therefore: E:[F]/RL:[OF]/RC:[C]. Table 7.4 shows summarized results of CVE-2015-0204.

Despite the fact that an exploitable functional code exists, a remediation action is available that shows an official fix to this vulnerability thereby lowering the base CVSS score from 8.3 to 6.9.

According to Redhat reports, Heartbleed vulnerability exploit code is known to exist and therefore the exploitability metric value is set to 'Functional' since it is in existence and operational for use. Security updates and patches for CVE-2014-0160 have been released by various vendors to protect their products from it, like Oracle for example, so Remediation

Listing 7.5: Computation of FREAK attack temporal score

```
1 TemporalScore = BaseScore * Exploitability * RemediationLevel
2                 * ReportConfidence
3               = 8.3 * 0.95 * 0.87 * 1
4               = 6.9
5
6 Thus CVE-2015-0204 Temporal score is 6.9
```

Table 7.5: CVE-2014-0160 Temporal Metric Values

Temporal Metric	Evaluation	Score
Exploitability (E)	Functional(F)	0.95
Remediation Level(RL)	Official-fix(OF)	0.87
Report Confidence(RC)	Confirmed (C)	1.00

Level for CVE-2014-0160 was 'Official-Fix'. The patch released also ensured that report confidence was turned to 'Confirmed'. The temporal vector for this vulnerability was, therefore: E:[F]/RL:[OF]/RC:[C]. Table 7.5 shows summarized results of CVE-2014-0160 Temporal metric values.

Applying the base metrics made the following adjustments to the Heartbleed base score:

The temporal score metrics for the Heartbleed vulnerability was adjusted from the base score to give it a temporal score of 8.3

7.4.3 Computation of Environmental Score

When CVE-IDs are added to the CVE dictionary they do not attach to it a specific environment to which they apply. Instead, upon returning to CVSS, a wide variety of environments in which applications of vulnerabilities an organization is exposed to is

Listing 7.6: Computation of Heartbleed attack temporal score

```
1 TemporalScore = BaseScore*Exploitability*RemediationLevel
2                 * ReportConfidence
3               = 10*0.95*0.87*1
4               = 8.3
5
6 Thus CVE-2014-0160's Temporal score is 8.3
```

Table 7.6: CVE-2015-0204 Environmental Metric Values

Environmental Metric	Evaluation	Score
Collateral Damage Potential(CDP)	None(N) - High(H)	0 - 0.5
Target Distribution(TD)	None(N) - High(H)	0 - 1.0
Confidentiality Req(CR)	Medium(M)	1.0
Integrity Req(IR)	Medium(M)	1.0
Availability Req(AR)	Medium(M)	1.0

given. Depending on the environment being studied the environmental score could vary. The algorithm for computing the environmental score is given in Listing 7.7.

For CVE-2015-0204 it was assumed that confidentiality, integrity, and availability are roughly equally important for the targeted systems. The temporal vector for this vulnerability is therefore: CDP:[N,H]/TD:[N,H]/ CR:[M]/ IR:[M]/AR:[M]. Table 7.6 shows all metrics required for this calculation.

Since confidentiality, integrity and availability requirements were evaluated to 'Medium' which gave them a score of 1.0 each, then the adjusted temporal score remains 6.8 as it has been computed earlier. If however any of these three had a value different from 1.0 then the study would have used the adjusted impact formula for the environmental score to make the necessary changes reflected in the new environment. The study could have also adjusted the base score because the base metric values accommodate the three variables as well. Thus the environmental value is computed as follows:

Assuming that confidentiality, integrity, and availability were roughly equally important for the targeted systems, and depending on the values for Collateral Damage Potential and Target Distribution, the environmental score could vary between 0.0 in case study one ('None', 'None') and 8.5 in case study two ('High', 'High'). Table 7.7 shows a scale that has a qualitative representation of CVSS score as proposed by the FIRST. Using this scale, it showed that CVE-2015-0204 is rated 'High' falling in the range of 7.0 - 8.9. This means that the potential impact that FREAK vulnerability can cause is rated high i.e. a successful exploit of CVE-2015-0204 had a significant physical or property damage or loss or even a significant loss of revenue or productivity.

Looking at Heartbleed's impact and knowing that it has the potential to cause DoS to end users of the targeted system, the study assumed that availability was more important than usual for the targeted systems. The temporal vector for this vulnerability was therefore: CDP:[N,H]/TD:[N,H]/ CR:[M]/ IR:[M]/AR:[H]. The environmental metrics for CVE-2014-0160 were summarised in Table 7.8.

Listing 7.7: Formual for computing environmental score

```
EnvironScore =((AdjustedTemporal + (10-AdjustedTemporal)
             *CollateralDamagePotential)*TargetDistribution)

AdjustedTemporal = TemporalScore recomputed with the
                   BaseScores Impact sub-equation
                   replaced with the AdjustedImpact
                   equation

AdjustedImpact = min(10,10.41*(1-(1-ConfImpact*Conf Req)
                *(1-IntegImpact*IntegReq)*(1-AvailImpact
                *AvailReq)))

CollateralDamagePotential = case CollateralDamagePotential of
                                 none:           0
                                 low:            0.1
                                 low-medium:     0.3
                                 medium-high:    0.4
                                 high:           0.5
                                 not defined:    0

TargetDistribution        = case TargetDistribution  of
                                 none:           0
                                 low:            0.25
                                 medium:         0.75
                                 high:           1.00
                                 not defined:    1.00

ConfReq            = case ConfReq of
                          low:            0.5
                          medium:         1.0
                          high:           1.51
                          not defined:    1.0

 IntegReq          = case IntegReq of
                          low:            0.5
                          medium:         1.0
                          high:           1.51
                          not defined:    1.0

AvailReq           = case AvailReq of
                          low:            0.5
                          medium:         1.0
                          high:           1.51
                          not defined:    1.0
Note that the final score ought to be rounded to one decimal
place
```

Listing 7.8: Computation of FREAK attack temporal score

```
AdjustedTemporal  = 6.9
EnvironScore = ((AdjustedTemporal+ (10-AdjustedTemporal)
               *CollateralDamagePotential)*TargetDistribution)
             = ((6.9+(10-6.9)*{0-0.5})*{0 - 1.0})

case 1: Collateral damage pontential = 0
                  Target Distribution = 0

          EnvironScore = round ((6.9+(3.2)*(0))*0)
                       =((6.9+(3.2)*0)*0)
                       = (6.9 +0)*0
                       = 6.8*0
                       = 0.0

case 2: Collateral damage pontential = 0.5
                  Target Distribution = 1

       EnvironScore = round ((6.9+(3.2)*(0.5))*1)
                       =((6.9+(3.2)*0.5)*1)
                       = (6.9 +1.6)*1
                       = 8.5*1
                       = 8.5

Thus CVE-2015-0204's Environmental score will fluctuate
between 0 and 8.5
```

Table 7.7: Qualitative Severity Rating Scale

Rating	CVSS Score
CVSS Score	0.0
Low	0.1 - 3.9
Medium	4.0 - 6.9
High	7.0 - 8.9
Critical	9.0 - 10.0

Table 7.8: CVE-2014-0160 Environmental Metric Values

Temporal Metric	Evaluation	Score
Collateral Damage Potential(CDP)	None(N) - High(H)	0 - 0.5
Target Distribution(TD)	None(N) - High(H)	0 - 1.0
Confidentiality Req(CR)	Medium(M)	1.0
Integrity Req(IR)	Medium(M)	1.0
Availability Req(AR)	High(H)	1.51

Working with our assumption that the targeted system's availability is more important than the temporal score, the Heartbleed vulnerability score will be adjusted by adding environmental metric values as shown in Listings 7.8 and 7.9.

Depending on the values for Collateral Damage Potential and Target Distribution, the environmental score could vary between 0.0 ('None', 'None') and 9.2 ('High', 'High').

7.5 Significance of the CVSS Scores to EAC

This section reflected on the significance of the computation made in this chapter for CVE-2015-0204 and CVE-2014-0160, thus the research looked at how each of the countries in EAC was fairing against these vulnerabilities. Tables 7.9 and 7.10 show vulnerability exposure rate for EAC countries to CVE-2015-0204 and CVE-2014-0160 in reference to datasets G and H from SHODAN. Table 7.9 shows 61.10% of the vulnerabilities data sets came from Kenya of which 77.67% of this is FREAK vulnerability (Table 7.10). With Tanzania coming second with 20.77% (of which 88.27% is FREAK vulnerability) then Uganda with 12.83% (of which 60.82% is FREAK vulnerability) and lastly Malawi with 5.30% (of which 84.62% is FREAK vulnerability).

Looking at CVE-2014-0160, it can be concluded that the rate at which Heartbleed affected EAC was not the same as that of FREAK vulnerability. Table 7.10 shows that of all the records of data that was collected as of end October 2016, 22.33% of Kenya's

Listing 7.9: Computation of Heartbleed attack environmental score

```
AdjustedImpact = min(10,10.41*(1-(1-ConfImpact*ConfReq)
                 *(1-IntegImpact*IntegReq)*(1-AvailImpact
                 *AvailReq)))
                = min(10,10.41*(1-(1-0.660*1)*(1  -0.660*1)
                  *(1-0.66*1.51))
                = min(10,10.41*(1-(1-0.660)*(1-0.660)
                  *(1-0.9966))
                = min(10,10.41*(1-(0.34*0.34*0.0034)
                = min(10,10.41*(1-0.00039304))
                = min(10.0,10.41*(0.99960696)
                = min (10.0,10.4)
                = 10.0

Adjusted impact takes the lowest value therefore our
adjusted impact is 10

AdjustedBase =(0.6*AdjustedImpact)+(0.4*Exploitability)
              1.5)*f(Impact)
             =((0.6*10)+(0.4*10.0)1.5)*1.1  76
             =10.0

AdjustedTemporal = AdjustedBase * Exploitability
                   * RemediationLevel * ReportConfidence
                 = 10.0*0.95*0.87*1.0
                 = 8.3

EnvironScore =((AdjustedTemporal+(10-AdjustedTemporal)
              *CollateralDamagePotential)*TargetDistribution)
             =((8.3+(10-8.3)*{0-0.5})*{0 - 1.0})
             =((8.3 + 1.7*{0 - 0.5})*{0 - 1.0}

case 1: Collateral damage pontential = 0
                Target Distribution = 0

          EnvScore = round ((8.3+(1.7)*(0))*0)
                   =((8.3+(3.2)*0)*0)
                   = (8.3 +0)*0
                   = 8.3*0
                   = 0.0

case 2: Collateral damage pontential = 0.5
                Target Distribution = 1

          EnvScore = round ((8.3 +(1.7)*(0.5))*1)
                   =((8.3+1.7*0.5)*1)
                   = (8.3 +0.85)*1
                   = 9.15*1
                   = 9.2
```

Table 7.9: Vulnerability Exposure Level in EAC

# of CVE-ID instances				
Country	**CVE-2015-0204**	**CVE-2014-0160**	#	%
Kenya	233	67	300	61.10
Tanzania	88	14	102	20.77
Uganda	38	25	63	12.83
Malawi	22	4	26	5.30
			491	100

Table 7.10: EAC Exposure to Vulnerabilities

	Country								
Vulnerability	**Kenya**	**%**	**Tanzania**	**%**	**Uganda**	**%**	**Malawi**	**%**	**#**
CVE-2015-0204	233	77.67	88	86.27	38	60.82	22	84.62	381
CVE-2014-0160	67	22.33	14	13.73	25	39.68	4	15.38	110
Total	300		102		63		26		491

computers were susceptible to Heartbleed vulnerability. While in Tanzania it was 14% of the computers that were prone to the Heartbleed, 39.68% for Uganda and 15.38% for Malawi. If an attacker is to target any of these countries Kenya for this vulnerability, based on Table 7.10, then Uganda will be highly affected than the rest of the countries with Tanzania being the least affected. The degree to which they are affected depends on the intention of the attacker.

What this means is that if an attacker is to explore these two vulnerabilities in each of these countries, a lot of damage or compromise would happen in Kenya because it has more devices exposed to the observed vulnerabilities than the other three countries with Malawi being the least country affected by it. One would see that this is a vulnerability that has the potential to cause a lot of harm to systems because of its high CVSS score. The damage caused ranges from compromising the confidentiality of data in the targeted systems to Integrity of the data which can be caused when an attacker alter user credentials or data content. In worst case scenario, the attacker may do both of the aforementioned examples coupled with denying end users services thus making the availability of systems void.

If all the vulnerabilities present at any given time are identified it is possible to gauge the levels of vulnerability for any given CVE-ID. These two vulnerabilities were just samples of the vulnerabilities found in EAC which also show how each of the vulnerabilities can be quantified so that we know to what extent the targeted country is susceptible to a specified vulnerability. Depending on the industry in which an organization is conducting its business, it would help them in prioritizing which vulnerability to focus on first based on how high the score is and how bad such a vulnerability can affect them.

7.6 Summary

This chapter focused on how vulnerabilities found in EAC can be quantified. It looked at two samples of the vulnerabilities that were found (CVE-2015-0204 and CVE-2014-0160) when data was collected in November 2016. The key value to quantifying vulnerabilities is the presence of CVE-ID along which are metrics attached to it which facilitate the computation. With all the metric values provided, it helps each of these countries to have a clear understanding of the vulnerabilities that they are facing and thus be well prepared to manage them before they get out of hand. Given the metrics that were collected this far for both vulnerabilities and that all the environmental assumptions hold then it follows that if Kenya was to be targeted, it would definitely want to mitigate these vulnerabilities because the environmental factors affect it more (up to 8.4) than the other three countries i.e it has more property to lose than the other three based on the data that was analysed.

Chapter 8

Conclusion

The research discussed within this document has covered several aspects all of which pertain to cybersecurity and how OSINT can be used to provide insight to some of the challenging issues that exist today in the cyber world. The key factor being mention in Section 1.1 was that OSINT has provided easy means through which cyber attackers find it easy to penetrate system than they did in the past two decades. It highlighted some of the jargon used in the cybersecurity circles in Chapter 2 which have been used throughout this study. Since EAC is in Africa, the study focused on seeing the perspective of African countries to cybersecurity by focusing on the basis for this view in Section 3.1 and factors that are driving cybersecurity risks in Section 3.2.

OSINT formed the basis of the data used for this study, with Chapter 4 explaining and expanding how IBR data was processed for use in Section 4.2 which was followed by SHODAN data in Section 4.4. Characteristics that were used for this study in IBR data sets and nine datasets from SHODAN (A, B, C, D, E, F, G, H and I) shown in Table 4.7 were explained in Sections 4.3 and 4.4 respectively.

This chapter offers a summary of the insights that have been acquired while conducting this study. In Section 8.1 a list of the key findings that were explained in Chapters 5, 6 and 7 are presented together with possible suggestion and derived lessons from the study. In section 8.2 an evaluation of the research goals to verify if all the objectives that were set in Section 1.2 have been met. In Section 8.3 an explanation of the impact of this study's findings is made and finally, it closes with ideas that point out possible future work that could extend from this study in section 8.4.

8.1 Key Findings

IBR and SHODAN datasets analyzed have exhibited variables that can prove to be very useful in the hands of cyber attackers should the attacker get hold of them. The ones discussed in this research are port numbers, their functionality and how they can be explored. It also looks at IP addresses, how network telescopes can be used as a tool to detect early malicious activities and how reuse of IP addresses can affect future IP address allocation. It also looked at various operating systems that SHODAN was able to pull banners from, ISPs, OpenSSL versions and its vulnerabilities. Products made by different vendors and how an attacker can get its details from them were also looked upon.

In addition to this, the research looked at device types that under normal circumstances are not supposed to be connected to the Internet and if they are found on it then proper measures need to be taken to ensure they do not leak information to attackers. Should proper measures fail on such devices, poor configurations may be detected by SHODAN as it has been shown in Tables 6.7 and 6.8 respectively. In addition, failure to change default passwords would provide easy access to attackers and give them unlimited access to such devices. Thus in this section, the study points towards drawing conclusions from what this research has explored based on the work covered from Chapters 4, 5, 6 and 7 respectively.

8.1.1 OpenSSL

In Section 6.6 this study looked at a cryptographic library used in many server products to secure communications against eavesdropping like MiTM attack. Despite the fact that its core functionality is to ensure safe and secure communications, SHODAN is able to identify the versions being used. Each of the versions being used comes with it vulnerabilities that may be unique to that version and such information is useful to an attacker. In most cases, the vendor provides patches for the vulnerabilities but these too need to be applied for them to take effect without which they are rendered useless. Table 6.9 provides a sample of the summarised results that show some of the OpenSSL versions that were detected by SHODAN and the type of vulnerability that is associated with it together with its CVE.

Using proven formula, the paper has shown how one can compute a score for such vulnerabilities in their specified environment using CVSS which got its values from CVE-IDs.

It also showed that two organizations may be exposed to the same vulnerability but if their environments are different then they would have different scores depending on how the CIA triad affects it. This was demonstrated Section 7.4.

8.1.2 Port and IP Addresses

In Sections 5.2 and 5.3, the study has shown that IBR targets ports and unallocated IP addresses. Some ports are used for internal network operations as such at no point are they supposed to receive traffic from an external network, for example, port 7/TCP an echo port which has a vulnerability for DoS threat as attackers use it to relay flooding data. The same goes for port 110/TCP for a local proxy which operates on a router and port 3389/TCP for remote desktop. All these ought to be closed for security purposes. Alternatively, in the case of port 3389/TCP, since it is used for a remote connection, an organization can reconfigure and use another port to serve the same purpose unlike it since it is now well known to act as a window to threats from outside the organization especially hackers.

Another area to look at is port 443/TCP which is used for HTTPS while port 80/TCP is for HTTP meaning it is more secure for browser communication to occur on 443/TCP than 80/TCP if certificates are set up properly. Proper measures need to be taken to ensure that ports 443/TCP (HTTPS), 80/TCP (HTTP), 21/TCP (FTP), 22/TCP (SSH), 23/TCP (TELNET), 161/UDP (SNMP) and 5060/UDP (SIP) are properly secured for fear of being used as a means of further access into the networks. Since certain ports cannot be closed due to the services that they run on them, it is proper and wise to ensure that all applications that use these ports are properly patched and kept up to date.

It is highly recommended that organizations have network telescopes in their networks to keep track of events happening in the unallocated IP address space of their companies as this could provide a breeding ground for future attackers as explained in explained in Section 5.4. It also explained the possibility of attackers to reuse the same IP address if it has not been allocated even after two or three years down the line. More has been explained in Section 5.4

8.1.3 Operating Systems

SHODAN has shown that other than Mac OS most of these commonly used operating systems were more likely to release their banners should a probe made by SHODAN be done to them. Details like IP addresses, MAC address and paths to certain software within serves are all acquired once SHODAN pulls banners from such operating systems. Ubuntu operating system dominates more than any other operating system in EAC, but the study and data sets did not provide more information as to why this is the case. Unix, Fedora, and ClearOS are the operating systems that appear the least in all four countries. A further study would help to clarify this scenario, otherwise, at the moment it is just a hypothesis.

The availability of these operating system was not uniform in all countries because some operating systems appeared in other countries and had zero attempts made in other countries to exploit a vulnerability. This does not offer any conclusive evidence that they do not exist, however it may be assumed that if they are there, then they are well patched and secure from SHODAN. Table 6.11 in Section 6.8 provides a summary of all the operating systems that were detected together with their statistics.

8.1.4 Internet Service Provider

In section 6.11 the study looked at ISP distribution in all four countries. Working with the results observed, it showed that the differences in ISP prevalence means that clients that were under the top ISPs are more prone to attack than the rest of the ISPs with low prevalence. Major ISPs need to take an active role in ensuring the safety and security of its clients as high prevalence of their cases could easily trigger keen interest to an attacker.

Each ISP has practices which it follows in order to run its businesses and deliver their service to its customers. Such principles need to have safety and security of its customers among its top priorities. Unfortunately, it looks like the bigger the ISP, the more its clients end up being picked by SHODAN. Meaning both big and small ISPs have the same problem to work with.

Looking at IBR data the same concept applied too in terms of IP address allocation. It would be a safety measure not to allocate large IP address blocks to companies whose growth rate is not well defined as this creates more breeding grounds for malicious activity which only cause more problems when unused IP addresses come into use

8.1.5 Device Type

In Section 6.12 an analysis of the device types was performed and results shown tabulated. In view of these outcomes, it became more clear why certain devices should never be connected to the Internet. Good examples are cameras and media devices. When devices like PBX and WAP are accommodated on user networks special precautions ought to be taken to ensure that they do not become listening devices for SHODAN as it can pull information from them which may be critical to the attacker and leave user systems and devices vulnerable.

As much as it becomes easier to monitor certain devices on the Internet like, printers for example, it also puts the device at risk of being accessed by remote attackers. Devices like printers and routers need to be properly secured and default passwords removed as an attacker can easily find his way around them and gain control of them. Looking at Malawi and Tanzania it showed how device types can increase over a period of a year. The actual number of these two countries over a 14 months period was very significant, which later led to new devices being detected. If this is the trend to go by then we should expect more devices should another analysis be carried out over six to twelve months from the last time data was collected.

The availability of device types found in these countries does not give a complete picture of how vulnerable these countries are but it gave room for consideration on how to safeguard them. For example, WAP devices have changed since 2015 and looking at the current trend on mobile banking applications that use them, then one should be worried of the future of mobile banking with such an increase in its applications. All devices with web interfaces need to be secured with strong passwords and remove all default passwords as SHODAN can easily detect them. Of the devices available, routers seem to be very prone to attacks hence the need to ensure that they are properly secured too. This study provides a starting point on what to look for because failure to know the device that is emitting the signal would make it harder to curb the problem.

8.1.6 Role of Vendors and their Products

One of the major items that were identified by SHODAN were products designed and programmed by various vendors across the globe. Each of the products identified had its own weakness otherwise it would not have appeared in the data set. A sample of these products and their vulnerabilities have been tabulated in Table 6.12 in Section 6.9. The

known vulnerabilities that were assigned CVE-IDs may have tested solutions and patches provided for but not all of its clients are up to date with such updates and patches. It is the role of the vendors to ensure that before such products are released to the market to ensure safe and secure coding practices were part and parcel of their guiding principles in developing them as this is the only safe way to keep their clients from harm in the cyber world. Attackers explore the vulnerabilities present in their products in order to gain access to systems thus if such vulnerabilities are minimised it gives clients a better chance of survival in the Internet world.

Vendors who develop devices like routers, switches, and printers among others need to ensure that they implement strategies that force their clients to change default passwords once they do an initial login. Alternatively, the credentials should be given a time frame to operate after which they expire. This will prevent easy access of credentials to attackers. The number of default passwords present in different sites on the Internet poses a huge threat to these devices. There is also need to minimise privileges that are attached to these default passwords as some of them give full control of devices making them dangerous should such passwords fall into the hands of the attacker. Full administrative rights of such devices only need to be activated once the default credentials are done away with.

8.2 Evaluation of Research Goals

There were four objectives that the study aimed to accomplish thus this section will be used as a yardstick to measure the success or failure of the objectives it was intended to achieve. The goals set in Section 1.2 are the ones being evaluated in this section and in the same order.

1. This research showed in Section 7.4, using two examples, how CVSS can be computed using CVE-IDs that were extracted from the SHODAN data sets G and H. The first one was the Heartbleed CVE and the second one was the FREAK vulnerability. Each of these two was exposed to different environments to see how they would all respond to such circumstances and gauge what impact it might have on its environment.

2. When looking at IBR in Sections 4.3, 5.3 and 5.4, the research showed that attackers can use either destination ports or IP source addresses to perform an attack which if not attended to may be reused yearly until later on move to the allocated IP address

space once it starts making random probes. The moment it finds one vulnerable client on the network it spreads throughout like a worm. DDoS is one the attacks that can be generated from IBR. This work has also shown that if source IP addresses in IBR remain unattended or are not monitored, they can repeatedly be used to perform malicious activities simply by changing the port of exploitation. Thus the importance of IBR cannot be overemphasized.

3. The aim was to have at least a one year gap in data collection between the initial data collection point and the next. Unfortunately, the data that was present from June 2015 was not accommodative of all the four countries. However, in Section 6.9, the study showed the changes that have occurred in Malawi and Tanzania for a period of 14 months (one year two months) by using three variables i.e. device type, operating systems, and ports.

4. In Section 6.8 this research identified devices that were present in the SHODAN data. It also identified security misconfigurations in the process in all the four countries. Apart from that, the study identified operating systems, products, OpenSSL, ports and ISPs as some of the variables that can be used to identify vulnerabilities in systems. In the case of OpenSSL and products, this research went further by identifying the type of attack that can occur and its associated CVE-ID.

Going back to the title which is the guiding principle of this research study, it is clear that OSINT data sets have some valued information that can be critical to safeguarding systems and applications should it be given enough attention and consideration. Having met the set goals this research shows potential to expand further on these findings as it may act as a benchmark for future studies.

8.3 Impact of the Findings

This study sets a benchmark for future work since work of similar nature has not been done in EAC and no papers have been published about it. In addition to this, it also helps our understanding of the current state of affairs in relation to cyber vulnerabilities. OSINT has been extensively used in other countries to track terrorists or project potential markets for businesses. The same tool can be used to enhance cybersecurity. With the way IBR operates, it has provided room for some anti-virus companies to predict future behavior of viruses and worms. The same can be said of cyber attackers whose current

cyber acts shows that companies need to deploy network telescopes in their network in order to prevent future attacks emanating from their dark space as it shows malicious activity before it moves to the allocated IP addresses.

Looking at this research's findings, organizations can take an active role in ensuring that certain ports for examples that have been neglected or are not being used ought to be attended to because they may not be as idle as they think. While they are not using them some devices outside their network could probably be listening to them. Knowing which operating systems are being targeted it would help companies and organization who are using them to be more cautious and careful on how they manage them should they be keen on continuing using or opt to other versions altogether.

The bottom line is there is more that we can learn through analysis of OSINT data and this research provides a glimpse of the things that may otherwise not be identified should OSINT be neglected and not taken seriously. It points users and clients in the direction which they need to focus on to keep their devices safe from cyber attacks. It helps readers to be aware of the potential threat that may be upon their door step that they would not have otherwise known and to those that already had an idea it provides as an alternative means for them to be informed. This applies a lot in instances where people overlook the notion of changing default passwords or carelessly leaving devices that are not supposed to be connected to the Internet like printers for example. It also helps readers of this research to be cautious of their system configurations as even they can be picked by SHODAN and later on used against them.

8.4 Future Work

Open source intelligence data will keep on growing in its use and, because of its evolving nature, more information will be added to the current version and provide more room for reference later on. It gives an opportunity to do a comparison of certain variables over time like in chapter 6 when comparing variables found in Malawi and Tanzania. Given the fact that this research addressed the principles behind the use of SHODAN and IBR data sets, a closer look at specific variable can be looked upon later on. Thus the possibility of future work based on what we have found cannot be overlooked.

1. Given the fact that CVE-IDs can be used to quantify vulnerabilities in various environments, a comparative analysis that involves quantifying the vulnerabilities

in each country can be carried out. This will look at specific products by different vendors and assess how each of these vulnerabilities is faring in different countries i.e. a specific list of products appearing in all the four countries will be compiled together with a list of accompanying CVE-IDs and compute their score.

2. An in-depth analysis using additional OSINT tools will be carried out to identify why certain operating systems are captured more by SHODAN than others. This will help in finding out the weaknesses in these operating systems so as to have them patched from such vulnerabilities.

3. Expanding on the scope to accommodate more countries and identify device types that are common in all the countries. This will be accompanied by assessing user behavior in handling such devices. For example, normal practices followed by bank clients who use mobile banking application on their phones.

4. Crosscheck on the progress of the variables used in this research to see if they have changed over time. This research showed that some variables have changed when we were looking at Malawi and Tanzania. This time around it will encompass all the variables that have been looked into. This will apply to both SHODAN and IBR.

5. Expand on the variables that were looked at from IBR data set as the research only looked at IP addresses and ports. The study covered on TCP port but never looked at UDP and ICMP ports. A geopolitical analysis will be carried out using the identifies coordinates that are part of the variables found in IBR data set.

6. Further analysis can be done to work out the compatibility of SHODAN and IBR data sets i.e. identifying variables that are common in both (ports for example) and identify which ones are similar or how the two data sets merge together to formulate a consolidated report.

These points can provide more insight into the things that were just looked at from a principle level in this paper. As such further work and analysis on this work cannot be overlooked. It will also provide more room to add in more OSINT sources and compare the findings in each of the cases when looking at similar variables thus giving us a new perspective altogether.

References

Abraham, A., Grosan, C., and Chen, Y. *Cyber Security and the Evolution of Intrusion Detection Systems. i-Manager's Journal on Future Engineering and Technology*, 1(1):74, 2005.

Abraham, A. and Thomas, J. *Distributed Intrusion Detection Systems: A Computational Intelligence Approach.* In *Applications of Information Systems to Homeland Security and Defense*, pages 107–137. IGI Global, 2006.

Abualola, H., Alhawai, H., Kadadha, M., Otrok, H., and Mourad, A. *An Android-based Trojan Spyware to Study the NotificationListener Service Vulnerability. Procedia Computer Science*, 83:465–471, 2016.

Acıiçmez, O. and Schindler, W. *A Vulnerability in RSA Implementations Due to Instruction Cache Analysis and Its Demonstration on OpenSSL.* In **Malkin, T.**, editor, *Topics in Cryptology – CT-RSA 2008: The Cryptographers' Track at the RSA Conference 2008, San Francisco, CA, USA, April 8-11, 2008. Proceedings*, pages 256–273. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. ISBN 978-3-540-79263-5.

Ahuja, S., Johari, R., and Khokhar, C. *EAST: Exploitation of Attacks and System Threats in Network.* In *Information Systems Design and Intelligent Applications*, pages 601–611. Springer, 2015.

Aiello, M., Avanzini, D., Chiarella, D., and Papaleo, G. *Worm Detection Using e-mail Data Mining.* pages 1–4. National Research Council, Institute IEIIT, Genoa, University of Genoa, Department of Computer and Information Sciences, Italy, 2006.

Aker, J. C. and Mbiti, I. M. *Mobile Phones and Economic Development in Africa. The Journal of Economic Perspectives*, 24(3):207–232, 2010.

Ali, A., Zavarsky, P., Lindskog, D., and Ruhl, R. *A software application to analyze the effects of temporal and environmental metrics on overall CVSS v2 score.* In *Internet Security (WorldCIS), 2011 World Congress on*, pages 109–113. IEEE, 2011a.

Ali, A. B. M., Shakhatreh, A. Y. I., Abdullah, M. S., and Alostad, J. *SQL-Injection Vulnerability Scanning Tool for Automatic Creation of SQL-Injection Attacks . Procedia Computer Science*, 3:453 – 458, 2011b. ISSN 1877-0509. World Conference on Information Technology.

Allen, M. *Social Engineering: A Means to Violate a Computer System. SANS Institute, InfoSec Reading Room*, 2006.

Alouneh, S., Kharbutli, M., and AlQurem, R. *Stack Memory Buffer Overflow Protection Based on Duplication and Randomization . Procedia Computer Science*, 21:250 – 256, 2013. ISSN 1877-0509.

Arce, I. and Levy, E. *An Analysis of the Slapper Worm. IEEE Security & Privacy*, 1(1):82–87, 2003.

Asokan, N., Niemi, V., and Nyberg, K. *Man-in-The-Middle in Tunnelled Authentication Protocols.* In *International Workshop on Security Protocols*, pages 28–41. Springer, 2003.

Backfried, G., Schmidt, C., Pfeiffer, M., Quirchmayr, G., Glanzer, M., and Rainer, K. *Open Source Intelligence in Disaster Management.* In *EISIC*, pages 254–258. 2012.

Bande, L. *The Making of Cybercrime Legislation in Malawi: A Comparative Analysis of Malawiś Proposed Cybercrime Law against International Standards.* 2013.

Bande, L. C. *A Case for Cybercrime Legislation in Malawi. Malawi LJ*, 5:93, 2011.

Banzhof, C., Cook, K., Helffrich, D., and Lawson, R. *Inventory Banagement-Based Computer Vulnerability Resolution System.* October 2004. US Patent App. 10/975,828.

Barford, P., Nowak, R., Willett, R., and Yegneswaran, V. *Toward a Model for Source Addresses of Internet Background Radiation.* In *Proceedings of the Passive and Active Measurement Conference*, pages 1–10. March 2006.

Baryamureeba, V. and Tushabe, F. *The Enhanced Digital Investigation Process Model.* In *Proceedings of the Fourth Digital Forensic Research Workshop*, pages 1–9. Citeseer, 2004.

Baskin, B. *Combating Spyware in the Enterprise.* Syngress Publishing, 2006.

Bayuk, J. L., Healey, J., Rohmeyer, P., Sachs, M. H., Schmidt, J., and Weiss, J. *Cyber Security Policy Guidebook.* John Wiley & Sons, 2012. ISBN 1118027809.

Bean, H. *The DNI's Open Source Center: An Organizational Communication Perspective. International Journal of Intelligence and Counterintelligence*, 20(2):240–257, 2007.

Berti, J. and Rogers, M. *Social Engineering: The Forgotten Risk. Information Security Management Handbook,*, 3:51–63, 2004.

Best, R. A. and Cumming, A. *Open source intelligence (osint): Issues for Congress.* volume 5. December, 2007.

Beurdouche, B., Bhargavan, K., Delignat-Lavaud, A., Fournet, C., Kohlweiss, M., Pironti, A., Strub, P.-Y., and Zinzindohoue, J. K. *A Messy State of the Union: Taming the Composite State Machines of TLS.* In *2015 IEEE Symposium on Security and Privacy*, pages 535–552. IEEE, 2015a.

Beurdouche, B., Delignat-Lavaud, A., Kobeissi, N., Pironti, A., and Bhargavan, K. *FLEXTLS: A Tool for Testing TLS Implementations.* In *9th USENIX Workshop on Offensive Technologies (WOOT 15)*. 2015b.

Bhargavan, K., Lavaud, A. D., Fournet, C., Pironti, A., and Strub, P. Y. *Triple handshakes and cookie cutters: Breaking and fixing authentication over TLS.* In *2014 IEEE Symposium on Security and Privacy*, pages 98–113. IEEE, 2014.

Blythe, S. E. *The Proposed Computer Laws of Uganda: Moving Toward Secure E-Commerce Transactions and Cyber-Crime Control. Journal of Management Policy and Practice*, 11(5):19–33, 2010.

Bowles, S. and Hernandez-Castro, J. *The first 10 years of the Trojan Horse Defence. Computer Fraud & Security*, 2015(1):5–13, 2015.

Bozorgi, M., Saul, L. K., Savage, S., and Voelker, G. M. *Beyond Heuristics: Learning to Classify Vulnerabilities and Predict Exploits.* In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 105–114. ACM, 2010.

Brenner, S. W. and Crescenzi, A. C. *State-Sponsored Crime: The Futility of the Economic Espionage Act. Houston Journal of International Law*, 28:389, 2006.

Bulgurcu, B., Cavusoglu, H., and Benbasat, I. *Information Security Policy Compliance: An Empirical Study of Rationality-Based Beliefs and Information Security Awareness. MIS Quarterly*, 34(3):523–548, 2010.

Burden, K. and Palmer, C. *Internet crime: Cyber Crime - A new Breed of Criminal? . Computer Law & Security Review*, 19(3):222 – 227, 2003. ISSN 0267-3649.

Burke, B. E. *The Evolution of Email Security: Symantec Brightmail Integrated Email Security Appliance. International Data Corporation*, January 2005.

Calof, J. L., Wright, S., and Fleisher, C. S. *Using Open Source Data in Developing Competitive and Marketing Intelligence. European journal of marketing*, 42(8):852–866, 2008.

Chariton, A. A., Degkleri, E., Papadopoulos, P., Ilia, P., and Markatos, E. P. *DCSP: Performant Certificate Revocation a DNS-Based Approach.* In *Proceedings of the 9th European Workshop on System Security*, pages 1 – 7. ACM, 2016.

Chen, Z., Zhang, Y., and Chen, Z. *A categorization Framework for Common Computer Vulnerabilities and Exposures. The Computer Journal*, 53(5):551–580, 2010.

Chien, E. *Anatomy of a SMSishing Attack.* July 2009. Date Accessed : 25 July 2016. URL `http://www.symantec.com/connect/blogs/anatomy-smsishing-attack`

Chindipha, S. D. and Irwin, B. *Cyber Vulnerability Assessment: Case Study of Malawi and Tanzania.* In **Van Niekerk, J. F.**, editor, *InProceedings of The African Cyber Citizenship Conference (ACCC)*, pages 105–121. Rhodes University, Nelson Mandela Metropolitan University, Nelson Mandela Metropolitan University PO Box 77000, Port Elizabeth, 6031, November 2015.
URL `http://accconference.nmmu.ac.za`

Choo, K.-K. R. *High Tech Criminal Threats to the National Information Infrastructure. Information Security Technical Report*, 15(3):104–111, 2010. ISSN 1363-4127. Computer Crime - A 2011 Update.

Choo, K.-K. R. *The Cyber Threat Landscape: Challenges and Future Research Directions . Computers & Security*, 30(8):719 – 731, 2011. ISSN 0167-4048.

Choo, K.-K. R., Smith, R. G., and McCusker, R. *Future Directions in Technology-Enabled Crime.* Australian Institute of Criminology Canberra, Australia, 2007.

Cole, K., Chetty, M., LaRosa, C., Rietta, F., Schmitt, D. K., Goodman, S. E., and Atlanta, G. *Cybersecurity in Africa: An assessment. Sam Nunn School of International Affairs*, 2008a.

Cole, K., Chetty, M., LaRosa, C., Rietta, F., Schmitt, D. K., Goodman, S. E., and Atlanta, G. *Cybersecurity in africa: An Assessment. Atlanta, Georgia, Sam Nunn School of International Affairs, Georgia Institute of Technology*, 2008b.

Cooke, E., Bailey, M., Mao, Z. M., Watson, D., Jahanian, F., and McPherson, D. *Toward Understanding Distributed Blackhole Placement.* In *Proceedings of the 2004 ACM Workshop on Rapid Malcode(WORM)*, pages 54–64. ACM, 2004.

Couture, E. *Web Application Injection Vulnerabilities: A Web App's Security Nemesis? SANS Institute Reading Room*, pages 1–34, May 2013.

Cowan, C., Wagle, F., Pu, C., Beattie, S., and Walpole, J. *Buffer Overflows: Attacks and Defenses for the Vulnerability of the Decade.* In *DARPA Information Survivability Conference and Exposition, 2000. DISCEX'00. Proceedings*, volume 2, pages 119–129. IEEE, 2000.

Cuijpers, C. *Legal Aspects of Open Source Intelligence : Results of the VIRTUOSO Project. Computer Law and Security Review*, 29(6):642 – 653, 2013. ISSN 0267-3649.

Cuppens, F., Cuppens-Boulahia, N., and Garcia-Alfaro, J. *Detection and Removal of Firewall Misconfiguration.* In *Proceedings of the 2005 IASTED International Conference on Communication, Network and Information Security*, volume 1, pages 154–162. 2005.

Das, T., Bhagwan, R., and Naldurg, P. *Baaz: A System for Detecting Access Control Misconfigurations.* In *USENIX Security Symposium*, pages 161–176. 2010.

Davis, D. *Understand the Levels of Privilege in the Cisco IOS.* June 2008. Date Accessed 22 November 2016.
URL `http://www.techrepublic.com/blog/data-center/understand-the-levels-of-privilege-in-the-cisco-ios-104552/`

Desouza, K. C. *Restructuring Government Intelligence Programs: A few Good Suggestions. Government Information Quarterly*, 22(3):342–353, 2005.

Dhamija, R., Tygar, J. D., and Hearst, M. *Why Phishing Works.* In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 581–590. ACM, 2006.

Difallah, D. E., Demartini, G., and Cudré-Mauroux, P. *Mechanical Cheat: Spamming Schemes and Adversarial Techniques on Crowdsourcing Platforms.* In *CrowdSearch*, pages 26–30. 2012.

Dobbins, R. *Mirai IoT Botnet Description and DDoS Attack Mitigation. Arbor Threat Intelligence*, 28, 2016.

D'Orazio, C. J. and Choo, K.-K. R. *A Technique to Circumvent SSL/TLS Validations on iOS Devices. Future Generation Computer Systems*, pages 1 – 9, August 2016. ISSN 0167-739X. doi:http://dx.doi.org/10.1016/j.future.2016.08.019.
URL `http://www.sciencedirect.com/science/article/pii/S0167739X16302801`

DuPaul, N. *SQL Injection Cheat Sheet & Tutorial: Vulnerabilities & How to Prevent SQL Injection Attacks.* January 2016. Date Accessed 28 July 2016.
URL `http://www.veracode.com/security/sql-injection`

Durumeric, Z., Kasten, J., Adrian, D., Halderman, J. A., Bailey, M., Li, F., Weaver, N., Amann, J., Beekman, J., Payer, M. *et al. The Matter of Heartbleed.* In *Proceedings of the 2014 Conference on Internet Measurement Conference*, pages 475–488. ACM, 2014.

Ellison, C. and Schneier, B. *Ten risks of PKI: What You're Not Being Told About Public Key Infrastructure. Computer Security Journal*, 16(1):1–7, 2000.

Eshete, B., Villafiorita, A., and Weldemariam, K. *Early Detection of Security Misconfiguration Vulnerabilities in Web Applications.* In *Availability, Reliability and Security (ARES), 2011 Sixth International Conference on*, pages 169–174. IEEE, 2011.

EventTracker. *Detecting and Patching FREAK Vulnerability (CVE- 2015-0204).* Whitepaper 7, EventTracke, EventTracker 8815 Centre Park Drive Columbia MD 21045, March 2015.
URL `https://www.eventtracker.com/wp-content/support-docs/How-To-Detecting-and-Patching-FREAK-Vulnerability.pdf`

Everett, C. *Ransomware: To Pay or not to Pay? Computer Fraud & Security*, 2016(4):8–12, 2016.

Fen, Y., Fuchao, Y., Xiaobing, S., Xinchun, Y., and Bing, M. *A New Data Randomization Method to Defend Buffer Overflow Attacks . Physics Procedia*, 24, Part C:1757 – 1764, 2012. ISSN 1875-3892. International Conference on Applied Physics and Industrial Engineering 2012.

Fette, I., Sadeh, N., and Tomasic, A. *Learning to Detect Phishing Emails.* In *Proceedings of the 16th International Conference on World Wide Web*, pages 649–656. ACM, 2007.

Flavián, C., Guinalíu, M., and Gurrea, R. *The Role Played by Perceived Usability, Satisfaction and Consumer Trust on Website Loyalty. Information & Management*, 43(1):1–14, 2006.

Fosnock, C. *Computer Worms: Past, Present, and Future. East Carolina University*, 8:1– 9, 2005.

Frank, R. *Managing Intellectual property. Journal of Accountancy*, 206(2):37, 2008.

Fruhwirth, C. and Mannisto, T. *Improving CVSS-Based Vulnerability Prioritization and Response with Context Information.* In *Proceedings of the 2009 3rd international Symposium on Empirical Software Engineering and Measurement*, pages 535–544. IEEE Computer Society, 2009.

Gates, C., Collins, M. P., Duggan, M., Kompanek, A., and Thomas, M. *More Netflow Tools for Performance and Security.* In *LISA*, volume 4, pages 121–132. 2004.

Gazette of the United Republic of Kenya, S. B. *The Cyber Security and Protection Bill.* Republic Of Kenya, Nairobi, senate bills no.12 edition, July 2016.

Gazette of the United Republic of Tanzania, B. S. *The Cybercrime Act.* The Republic of Tanzania, April 2015.

Gereda, S. L. *The Electronic Communications and Transactions Act. Telecommunications Law in South Africa*, pages 262 – 294, 2006.
URL `http://thornton.co.za/resources/telelaw12.pdf`

Ghafoor, I., Jattala, I., Durrani, S., and Tahir, C. M. *Analysis of OpenSSL Heartbleed Vulnerability for Embedded Systems.* In *Multi-Topic Conference (INMIC), 2014 IEEE 17th International*, pages 314–319. Dec 2014. doi:10.1109/INMIC.2014.7097358.

Gibson, S. *Open Source Intelligence: An Intelligence Lifeline. The RUSI Journal*, 149(1):16–22, 2004.

Glassman, M. and Kang, M. J. *Intelligence in the Internet Age: The Emergence and Evolution of Open Source Intelligence (OSINT). Computers in Human Behavior*, 28(2):673 – 682, 2012. ISSN 0747-5632.

Goitom, H. *Crime and Law Enforcement, Cybercrime, Freedom of Speech.* June 2015.
URL `http://www.loc.gov/law/foreign-news/article/tanzania-cybercrimes-bill-enacted/`

Gorbenko, A., Kharchenko, V., Tarasyuk, O., and Romanovsky, A. *Intrusion-Avoidance Via System Diversity. Information & Security: An International Journal*, 28(1):154–158, 2012.

Gorman, S. P., Schintler, L., Kulkarni, R., and Stough, R. *The Revenge Of Distance: Vulnerability Analysis of Critical Information Infrastructure. Journal of Contingencies and Crisis Management*, 12(2):48–63, 2004.

Greenberg, A. and Chamberlain, M. *This Machine Kills Secrets.* Tantor Media, Incorporated, 2012.

Gutmann, P. *Encrypt-then-MAC for Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS).* 2014.

Hadnagy, C. *Social Engineering: The art of Human Hacking.* John Wiley & Sons, 2010.

Hansen, L. and Nissenbaum, H. *Digital Disaster, Cyber Security, and the Copenhagen School. International Studies Quarterly*, 53(4):1155–1175, 2009.

Hariri, S., Qu, G., Dharmagadda, T., and Ramkishore, M. *Vulnerability Analysis of Faults and Attacks in Large-Scale Networks. IEEE Security and Privacy magazine*, pages 49–54, October and November 2003.

Harris, A., Goodman, S., and Traynor, P. *Privacy and Security Concerns Associated with Mobile Money Applications in Africa. Wash. JL Tech. & Arts*, 8:245, 2012.

Hayashi, K. *W32.hllw.moega.* 2004. Date Accessed 10 June 2016.
URL `https://www.symantec.com/security_response/writeup.jsp?docid=2003-080813-3234-99`

Heartbleed. *The Heartbleed Bug.* April 2014. Date Accessed: 30 October 2016.
URL `http://heartbleed.com/`

Hill, K. *The Crazy Things a Savvy Shodan Searcher Can Find Exposed on the Internet.* May 2013. Date Accessed: 07/August/2015.
URL `http://www.forbes.com/sites/kashmirhill/2013/09/05/the-crazy-things-a-savvy-shodan-searcher-can-find-exposed-on-the-internet/`

Hoffman, P. *SMTP Service extension for Secure SMTP Over Transport Layer Security.* 2002.

Holm, H. and Afridi, K. K. *An expert-based investigation of the Common Vulnerability Scoring System . Computers & Security*, 53:18 – 30, 2015. ISSN 0167-4048.

Houmb, S. H., Franqueira, V. N., and Engum, E. A. *Quantifying Security Risk Level from CVSS Estimates of Frequency and Impact. Journal of Systems and Software*, 83(9):1622–1634, 2010.

Hunter, S. O., Irwin, B., and Stalmans, E. *Real-time Distributed Malicious Traffic Monitoring for Honeypots and Network Telescopes.* In *Information Security for South Africa, 2013*, pages 1–9. IEEE, 2013.

Hypponen, M. *Malware Goes Mobile. Scientific American*, 295(5):70–77, 2006.

IBM. *An overview of the SSL or TLS handshake.* Technical report, IBM Knowledge Center, 2017. Date Accessed 31 October 2017.

Ingols, K., Chu, M., Lippmann, R., Webster, S., and Boyer, S. *Modeling Modern Network Attacks and Countermeasures Using Attack Graphs.* In *Computer Security Applications Conference, 2009. ACSAC'09. Annual*, pages 117–126. IEEE, 2009.

Irani, D., Balduzzi, M., Balzarotti, D., Kirda, E., and Pu, C. *Reverse Social Engineering Attacks in Online Social Networks.* In *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*, pages 55–74. Springer, 2011.

Irwin, B. *A network Telescope Perspective of the Conficker Outbreak.* In *2012 Information Security for South Africa*, pages 1–8. IEEE, 2012.

Irwin, B. V. W. *A framework for the application of network telescope sensors in a global IP network.* Ph.D. thesis, Rhodes University, 2011.

Ivaturi, K. and Janczewski, L. *A Taxonomy for Social Engineering Attacks.* In *International Conference on Information Resources Management, Centre for Information Technology, Organizations, and People (June 2011)*. 2011.

Jackon, D. *New Man-in-the-Middle Attacks Leveraging Rogue DNS.* March 2014. Date Accessed 15 November 2016.
URL `https://info.phishlabs.com/blog/new-man-in-the-middle-attacks-leveraging-rogu`

Jagatic, T. N., Johnson, N. A., Jakobsson, M., and Menczer, F. *Social Phishing. Communications of the ACM*, 50(10):94–100, 2007.

Jang, Y.-S. and Choi, J.-Y. *Detecting {SQL} Injection Attacks Using Query Result Size . Computers & Security*, 44:104 – 118, 2014. ISSN 0167-4048.

Jelic, F. *Man In The Middle Attacks.* October 2016. Date Accessed 15 November 2016.
URL `https://www.deepdotweb.com/2016/10/10/man-in-the-middle-attacks/`

Jensen, M. *ICT in Africa. World*, 2:1–15, February 2015.

John, J. P., Moshchuk, A., Gribble, S. D., and Krishnamurthy, A. *Studying Spamming Botnets Using Botlab.* In *NSDI*, volume 9, pages 291–306. 2009.

Joshi, Y. and Singh, A. *A study on Cyber Crime and Security Scenario in India. International Journal of Engineering and Management Research*, 3(3):13–18, 2013.

Judge, K. *OpenSSL Vulnerability Could be Exploited for Man-in-the-Middle Attacks.* June 2014. Date Accessed 1 November 2016.
URL `https://blog.comodo.com/it-security/openssl-vulnerability-exploited-man-middl`

Kim, Y., Lee, J., Han, H., and Choe, K.-M. *Filtering False Alarms of Buffer Overflow Analysis Using {SMT} Solvers . Information and Software Technology*, 52(2):210 – 219, 2010. ISSN 0950-5849.

Kristensen, T. *The Big Picture on Big Flaws: RPC DCOM Vulnerability - What went wrong. Network Security*, 2003(9):19 – 20, 2003. ISSN 1353-4858.
URL `http://www.sciencedirect.com/science/article/pii/S1353485803009115`

Krombholz, K., Hobel, H., Huber, M., and Weippl, E. *Social Engineering Attacks on the Knowledge Worker.* In *Proceedings of the 6th International Conference on Security of Information and Networks*, SIN '13, pages 28–35. ACM, New York, NY, USA, 2013. ISBN 978-1-4503-2498-4.

Laqueur, W. *Postmodern Terrorism. Foreign Affairs*, 75(5):24–36, September – October 1996.
URL `http://www.jstor.org/stable/20047741`

Lawack, V. A. *Mobile Money, Financial Inclusion and Financial Integrity: The South African Case. Wash. JL Tech. & Arts*, 8:317, 2012.

Lee, C., Yi, L., Tan, L.-H., Goh, W., Lee, B.-S., and Yeo, C.-K. *A Wavelet Entropy-Based Change Point Detection on Network Traffic: A Case Study of Heartbleed Vulnerability.* In *Cloud Computing Technology and Science (CloudCom), 2014 IEEE 6th International Conference on*, pages 995–1000. IEEE, 2014.

Lenstra, A., Hughes, J. P., Augier, M., Bos, J. W., Kleinjung, T., and Wachter, C. *Ron was Wrong, Whit is Right.* Technical report, International Association for Cryptologic Research (IACR), 2012.

Lichtman, D. and Posner, E. *Holding Internet Service Providers Accountable. Supreme Court Economic Review*, pages 221–259, 2006.

Liganga, F. B. *An Assessment of the Legal System in Relation to the Increasing Rates of Cyber Crimes in Tanzania . Liganga, Fikiri B., An Assessment of the Legal System in Relation to the Increasing Rates of Cyber Crimes in Tanzania, Forthcoming*, 2012.

Limbu, Y. B., Wolf, M., and Lunsford, D. L. *Consumers' Perceptions of Online Ethics and its Effects on Satisfaction and Loyalty. Journal of Research in Interactive Marketing*, 5(1):71–89, 2011.

Lindorfer, M., Di Federico, A., Maggi, F., Comparetti, P. M., and Zanero, S. *Lines of Malicious Code: Insights into the Malicious Software Industry.* In *Proceedings of the 28th Annual Computer Security Applications Conference*, pages 349–358. ACM, 2012.

Liu, Y., Tome, W., Zhang, L., Choffnes, D., Levin, D., Maggs, B., Mislove, A., Schulman, A., and Wilson, C. *An End-to-End Measurement of Certificate Revocation in the Web's PKI.* In *Proceedings of the 2015 ACM Conference on Internet Measurement Conference*, pages 183–196. ACM, 2015.

Longe, O., Ngwa, O., Wada, F., Mbarika, V., and Kvasny, L. *Criminal Uses of Information & Communication Technologies in Sub-Saharan Africa: Trends, Concerns and Perspectives. Journal of Information Technology Impact*, 9(3):155–172, 2009.

Maddison, M. *Risk Angles - Five Questions on the Evolution of Cyber Security.* 2013.

Magee, M. *W32.Korgo.F.* February 2007. Date Accessed: 9 June 2016. URL `https://www.symantec.com/security_response/writeup.jsp?docid=2004-060111-5322-99`

Magutu, P. O., Ondimu, G. M., and Ipu, C. J. *Effects of Cybercrime on State Security: Types, Impact and Mitigations with the Fiber Optic Deployment in Kenya. Journal of Information Assurance & Cybersecurity*, 2011(1):1–20, 2011.

Makulilo, A. B. *Cyber law in kenya. Information & Communications Technology Law*, 22(1):86–87, 2013.

Mambi, A. J. *ICT Law Book: A Source Book for Information and Communication Technologies and Cyber Law in Tanzania & East African Community.* African Books Collective, 2010.

McCrohan, K. F., Engel, K., and Harvey, J. W. *Influence of Awareness and Training on Cyber Security. Journal of internet Commerce*, 9(1):23–41, 2010.

Mell, P., Kent, K. A., and Romanosky, S. *The Common Vulnerability Scoring System (CVSS) and its Applicability to Federal Agency Systems*. Citeseer, 2007.

Mell, P., Scarfone, K., and Romanosky, S. *Common Vulnerability Scoring System. IEEE Security & Privacy*, 4(6):85–89, 2006.

Microsoft. *Buffer Overrun In RPCSS Service Could Allow Code Execution*. September 2003. Date Accessed: 8 June 2016.
URL `http://www.microsoft.com/technet/security/bulletin/MS03-039.mspx`

MIcrosoft. *Microsoft Security Bulletin MS12-020 - Critical*. March 2012. Date Accessed: 9 June 2016.
URL `https://technet.microsoft.com/en-us/library/security/ms12-020.aspx`

Mitre. *CVE-1999-0680: Windows NT Terminal Server Vulnerability*. 1999. Date Accessed: 9 June 2016.
URL `http://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-1999-0680`

Mitre. *CVE-2008-0852: Unspecified vulnerability in Freesshd 1.2*. 2008. Date Accessed: 9 June 2016.
URL `http://cve.mitre.org/cgi-bin/cvename.cgi?name=2008-0852`

Mitre. *CVE-2010-1099: Integer overflow in Apple Safari*. 2010a. Date Accessed: 9 June 2016.
URL `https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2010-1099`

Mitre. *CVE-2010-1099: Integer overflow in Apple Safari*. 2010b. Date Accessed: 9 June 2016.
URL `https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2010-1099`

Mitre. *CVE-2010-1100: Integer overflow in Arora*. 2010c. Date Accesses: 9 June 2016.
URL `http://cve.mitre.org/cgi-bin/cvename.cgi?name=2010-1100`

Mitre. *CVE-2010-1101: Integer overflow in Alexander Clauss iCab*. 2010d. Date Accessed: 9 June 2016.
URL `https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2010-1101`

Mitre. *CVE-2010-1102: Integer overflow in OmniWeb*. 2010e. Date Accessed: 9 June 2016.
URL `https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2010-201102`

Mitre. *CVE-2010-1103: Integer overflow in Stainless*. 2010f. Date Accessed: 9 June 2016.
URL `http://cve.mitre.org/cgi-bin/cvename.cgi?name=2010-1103`

Mitre. *Emerson DeltaV CVE-2012-4703 Denial of Service Vulnerability*. 2010g. Date Accesses: 9 June 2016.
URL `http://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2012-4703`

Mitre. *CVE-2012-4702: 360 Systems Maxx, Image Server Maxx, and Image Server 2000 Viulnerability*. 2012a. Date Accessed: 9 June 2016.
URL `http://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2012-4702`

Mitre. *CVE-2012-5345: Buffer Overflow in the Remote Command Server*. 2012b. Date Accessed: 9 June 2016.
URL `http://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2012-5345`

Mitre. *Microsoft Remote Desktop Protocol CVE-2012-0002 Remote Code Execution Vulnerability*. 2012c. Date Accessed: 9 June 2016.
URL `http://www.cve.mitre.org/cgi-bin/cvename.cgi?name=cve-2012-0002`

Mitre. *CVE - Common Vulnerabilities and Exposures[CVE]*. 2015. Accessed on 12 May 2015.
URL `http://cve.mitre.org/`

Mitre. *Common Vulnerabilities Exposures: The Standard for Information Security Vulnerability Names*. August 2016. Date Accessed : 25 August 2016.
URL `https://cve.mitre.org/about/`

Moore, D., Shannon, C., Voelker, G. M., and Savage, S. *Network Telescopes: Technical Report*. Department of Computer Science and Engineering, University of California, San Diego, July 2004.

Moore, D., Shannon, C. *et al. Code-Red: a Case Study on the Spread and Victims of an Internet Worm*. In *Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurment*, pages 273–284. ACM, 2002.

Mpofu, T. P., Elisa, N., and Gati, N. *The Heartbleed Bug: An Open Secure Sockets Layer Vulnerability. International Journal of Science and Research (IJSR)*, pages 1470 – 1473, 2012.

Mukkamala, S., Sung, A., and Abraham, A. *Cyber Security Challenges: Designing Efficient Intrusion Detection Systems and Antivirus Tools. Vemuri, V. Rao, Enhancing Computer Security with Smart Technology.(Auerbach, 2006)*, pages 125–163, 2005.

Murphy, D. R. and Murphy, R. H. *Teaching Cybersecurity: Protecting the Business Environment.* In *Proceedings of the 2013 on InfoSecCD '13: Information Security Curriculum Development Conference*, pages 88–93. ACM, New York, NY, USA, 2013. ISBN 978-1-4503-2547-9.

Nachenberg, C. *Computer Virus-Coevolution. Communications of the ACM*, 50(1):46–51, 1997.

Nagpal, N. B. and Nagpal, B. *Preventive Measures for Securing Web Applications Using Broken Authentication and Session Management Attacks: A Study.* In *International Conference on Advances in Computer Engineering and Applications (ICACEA)*, volume 24. October 2014.

Nam, S. Y., Kim, D., Kim, J. *et al. Enhanced ARP: Preventing ARP Poisoning-Based Man-in-The-Middle Attacks. IEEE communications letters*, 14(2):187–189, 2010.

Nichols, E. A. and Peterson, G. *A Metrics Framework to Drive Application Security Improvement. IEEE Security & Privacy*, 5(2):88–91, 2007.

Nicholson, A., Webber, S., Dyer, S., Patel, T., and Janicke, H. *Scada security in the light of cyber-warfare. Computers & Security*, 31(4):418–436, 2012.

Nykodym, N., Taylor, R., and Vilela, J. *Criminal Profiling and Insider Cyber Crime . Computer Law & Security Review*, 21(5):408 – 414, 2005. ISSN 0267-3649.

Ollmann, G. *Botnet Communication Topologies. Damballa*, 30:2009, September 2009a.
URL `https://www.damballa.com/`

Ollmann, G. *Extracting C&C from Malware : The Role of Malware Sample Analysis in Botnet Detection . Damballa*, 2009b.
URL `http://www.damballa.comÂ\protect\unhbox\voidb@x\penalty\@M\`

O'Mahony, J. *Stuxnet-Worm- Increased-Irans-Nuclear-Potential.* 2015. Date Accessed 15 February 2015.
URL `http://www.telegraph.co.uk/technology/news/10058546/Stuxnet-worm-increased-Irans-nuclear-potential.html`

Otey, M. *SQL Server TCP and UDP Ports.* April 2014. Date Accessed 5 October 2016. URL `http://sqlmag.com/sql-server/sql-server-tcp-and-udp-ports`

Pallaris, C. *Open Source Intelligence: A Strategic Enabler of National Security. Center for Security Studies*, 3(32), April 2008.

Pang, R., Yegneswaran, V., Barford, P., Paxson, V., and Peterson, L. *Characteristics of Internet Background Radiation.* In *Proceedings of the 4th ACM SIGCOMM conference on Internet measurement*, pages 27–40. ACM, October 2004.

Pannu, G. K. *A Survey on Web Application Attacks. IJCSIT) International Journal of Computer Science and Information Technologies*, 5(3):4162–4166, 2014. ISSN:0975-9646.

Petajasoja, S., Kortti, H., Takanen, A., and Tirila, J.-M. *IMS Threat and Attack Surface Analysis Using Common Vulnerability Scoring System.* In *Computer Software and Applications Conference Workshops (COMPSACW), 2011 IEEE 35th Annual*, pages 68–73. IEEE, 2011.

Piccard, P. and Faircloth, J. *Combating Spyware in the Enterprise.* 2006.

Polakis, I., Kontaxis, G., Ioannidis, S., and Markatos, E. P. *Dynamic Monitoring of Dark IP Address Space (Poster).* In *International Workshop on Traffic Monitoring and Analysis*, pages 193–196. Berlin Heidelberg Springer, April 2011.

Ponder-Sutton, A. M. *Chapter 1 - The Automating of Open Source Intelligence* . In **Layton, R. and Watters, P. A.**, editors, *Automating Open Source Intelligence*, pages 1 – 20. Syngress, Boston, 2016. ISBN 978-0-12-802916-9.

Porras, P. A., Saidi, H., and Yegneswaran, V. *A Foray into Conficker's Logic and Rendezvous Points.* In *LEET*. 2009.

Postel, J. and Reynold, J. *Telnet Protocol Specification.* 1983. Date Accessed: 9 June 2016.
URL `https://tools.ietf.org/html/rfc854`

Pringle, R. W. *The Limits Of OSINT: Diagnosing The Soviet Media, 1985-1989. International Journal of Intelligence and CounterIntelligence*, 16(2):280–289, 2003.

Provos, N. *et al. A Virtual Honeypot Framework.* In *USENIX Security Symposium*, volume 173, pages 1–14. 2004.

Pu, Y., Chen, X., Cui, X., Shi, J., Guo, L., and Qi, C. *Data Stolen Trojan Detection Based on Network Behaviors. Procedia Computer Science*, 17:828–835, 2013.

Quach, T. N., Thaichon, P., and Jebarajakirthy, C. *Internet Service Providers' Service Quality and its Effect on Customer Loyalty of Different Usage Patterns. Journal of Retailing and Consumer Services*, 29:104 – 113, 2016. ISSN 0969-6989.

Rainie, L., Duggan, M., and Tyson, A. *Heartbleed's Impact.* Report, Pew Research Center, April 2014.

Red-Hat. *FREAK: OpenSSL vulnerability (CVE-2015-0204)* . May 2015. Date Accessed 10 November 2016.
URL `https://access.redhat.com/articles/1369543`

RedHat. *CVE-2015-0204*. January 2015a.
URL `https://access.redhat.com/security/cve/cve-2015-0204`

RedHat. *Factoring RSA Export Keys - FREAK (CVE-2015-0204)*. March 2015b. Date Accessed: 26 October 2016.
URL `https://access.redhat.com/blogs/766093/posts/1976563`

Riccardo, P. and Luca, V. *3AKEP: Triple-Authenticated Key Exchange Protocol for Peer-to-Peer VoIP Applications . Computer Communications*, 85:28 – 40, 2016. ISSN 0140-3664.
URL `http://www.sciencedirect.com/science/article/pii/S0140366416301347`

Robbins, A. *Effective Awk Programming: Universal Text Processing and Pattern Matching.* " O'Reilly Media, Inc.", 2015.

Roman, R., Alcaraz, C., and Lopez, J. *The Role of Wireless Sensor Networks in the Area of Critical Information Infrastructure Protection. Information Security Technical Report*, 12(1):24–31, 2007.

Rosenquist, M. *The Future Evolution of Cybersecurity*. Intel Corp, Technology, October 2014. Cybersecurity Prediction Conference.

Rowe, B., Reeves, D., and Gallaher, M. *The Role of Internet Service Providers in Cyber Security.* Institute for Homeland Security Solutions, 2009.

Royal, P. *Analysis of the Kraken Botnet. Damballa, Apr*, 9, April 2008.

Saini, H., Rao, Y. S., and Panda, T. *A review of Cyber Crimes and their Impacts. International Journal of Engineering Research and Applications*, 2(2):202–209, 2012.

Savage, K., Coogan, P., and Lau, H. *The evolution of ransomware.* 2015.

Scarfone, K. and Mell, P. *An Analysis of CVSS Version 2 Vulnerability Scoring.* In *Proceedings of the 2009 3rd International Symposium on Empirical Software Engineering and Measurement*, pages 516–525. IEEE Computer Society, 2009.

Schaurer, F. and Störger, J. *Guide to the Study of Intelligence. The Evolution of the Open Source Intelligence (OSINT). The Intelligencer, Association of Former Intelligence Officers*, 2, 2011.

Schearer, M. *SHODAN for Penetration Testers.* 2010.

Schiffman, M. *A Complete Guide to The Common Vulnerability Scoring System (CVSS). Cisco CIAG*, June 2005.
URL `http://packetfactory.openwall.net/papers/CVSS/guide/index.html`

Schiffman, M., Wright, A., Ahmad, D., and Eschelbeck, G. *The Common Vulnerability Scoring System. National Infrastructure Advisory Council, Vulnerability Disclosure Working Group, Vulnerability Scoring Subgroup*, 2004.

Schultz, E. *The MSBlaster Worm: Going from Bad to Worse. Network Security*, 2003(10):4–8, 2003.

Schultz, E. *RPC in Windows Systems: What You Don't Know Could Hurt You. Network Security*, 2004(6):5–8, 2004.

Shannon, C. and Moore, D. *The Spread of the Witty Worm. IEEE Security & Privacy*, 2(4):46–50, 2004.

Shannon, H. *W32.Sasser.B.Worm.* February 2007.
URL `https://www.symantec.com/security_response/writeup.jsp?docid=2004-050114-1001-99`

Shar, L. K. and Tan, H. B. K. *Predicting {SQL} Injection and Cross Site Scripting Vulnerabilities Through Mining Input Sanitization Patterns. Information and Software Technology*, 55(10):1767 – 1780, 2013. ISSN 0950-5849.

Sheldon, J. *State of the Art: Attackers and Targets in Cyberspace. Journal of Military and Strategic Studies*, 14(2), 2012.

Singh, J., Kumar, K., Sachdeva, M., and Sidhu, N. *DDoS Attacks Simulation using Legitimate and Attack Real Data Sets. International Journal of Scientific & Engineering Research*, 3(6):1–5, 2012.

Singh, K. and Sharma, S. *Combating Broken Authentication and Session Management Attacks. IJCAT - International Journal of Computing and Technology*, 2(5):1–6, May 2015. ISSN : 2348 - 6090.

Smith, M. A. and Kollock, P. *Communities in Cyberspace.* Psychology Press, 1999. ISBN 0415191408.

Speedguide. *Port 445 details.* 2015. Date Accesses: 9 June 2016. URL `http://www.speedguide.net/port.php?port=445`

Speedguide. *Ports Database.* 2016a. Date Accessed: 5 October 2016. URL `http://www.speedguide.net/port.php?port=1433`

Speedguide. *Ports Database.* 2016b. Date Accessed: 5 October 2016. URL `http://www.speedguide.net/port.php?port=12203`

Speedguide. *Ports Database.* 2016c. Date Accessed: 5 October 2016. URL `http://www.speedguide.net/port.php?port=80`

Speedguide. *Ports Database.* 2016d. Date Accessed: 5 October 2016. URL `http://www.speedguide.net/port.php?port=139`

Speedguide. *Ports Database.* 2016e. Date Accessed: 9 June 2016. URL `http://www.speedguide.net/ports.php?filter=risk`

Spitzner, L. *Honeypots: Tracking Hackers*, volume 1. Addison-Wesley Reading, 2003.

Srinivasan, R. *Rpc: Remote procedure call protocol specification version 2.* 1995.

Steele, R. *Open Source Intelligence. Handbook of intelligence studies*, pages 129–147, 2007.

Stillions, T. *Spawar expert discusses getting ahead of the growing cyber threat. Space and Naval Warfare Systems Command*, 13:2, 2012.

Stone-Gross, B., Cova, M., Cavallaro, L., Gilbert, B., Szydlowski, M., Kemmerer, R., Kruegel, C., and Vigna, G. *Your Botnet is my Botnet: Analysis of a Botnet Takeover.* In *Proceedings of the 16th ACM conference on Computer and communications security*, pages 635–647. ACM, 2009.

Swart, I., Irwin, B., and Grobler, M. *Towards a Platform to Visualize the State of South Africa's Information Security.* In *Information Security for South Africa (ISSA), 2014*, pages 1–8. IEEE, 2014.

Swart, I. P. *Pro-active Visualisation of Cyber Security on a National Level: a South African Case Study.* Phd thesis, Rhodes University, 2015.

Symantec. *State of Enterprise Security 2010.* 2010. (Date Accessed 2 September 2016).

Taylor, P. A. *From Hackers to Hacktivists: Speed Bumps on the Global Superhighway? New Media & Society*, 7(5):625–646, 2005.

Tetri, P. and Vuorinen, J. *Dissecting Social Engineering. Behaviour & Information Technology*, 32(10):1014–1023, 2013.

Thaichon, P., Lobo, A., Prentice, C., and Quach, T. N. *The Development of Service Quality Dimensions for Internet Service Providers: Retaining Customers of Different Usage Patterns . Journal of Retailing and Consumer Services*, 21(6):1047 – 1058, 2014. ISSN 0969-6989. doi:http://dx.doi.org/10.1016/j.jretconser.2014.06.006. URL `http://www.sciencedirect.com/science/article/pii/S0969698914000836`

Thomson, K. and von Solms, R. *Using Knowledge Creation and Agency Theory to Shape an Information Security Obedient Culture.* In *Proceedings of the 11th IFIP TC 11.1 Working Conference on Information Security Management*, page 14. 2008.

Thomson, K.-L., von Solms, R., and Louw, L. *Cultivating an Organizational Information Security Culture. Computer Fraud & Security*, 2006(10):7–11, 2006.

Topalovic, E., Saeta, B., Huang, L.-S., Jackson, C., and Boneh, D. *Towards Short Lived Certificates. Web 2.0 Security and Privacy*, 2012.

Tsalis, N. and Gritzalis, D. *Hacking Web Intelligence: Open Source Intelligence and Web Reconnaissance Concepts and Techniques. Computers & Security*, 55:113 –, 2015. ISSN 0167-4048.

Tushabe, F. and Baryamureeba, V. *Cyber Crime in Uganda: Myth or Reality?* In *Proceedings of World Academy of Science, Engineering and Technology*, volume 8, pages 66–70. 2005.

US-CERT. *Alert (TA14-098A): OpenSSL 'Heartbleed' vulnerability (CVE-2014-0160).* April 2014. Date Acessed 2 November 2016.
URL `https://www.us-cert.gov/ncas/alerts/TA14-098A`

Vadehra, R., Chowdhary, N., and Malhotra, J. *Impact Evaluation of Distributed Denial of Service Attacks using NS2. International Journal of Security and Its Applications*, 9(8):303–316, 2015.
URL `http://dx.doi.org/10.14257/ijsia.2015.9.8.27`

Van Niekerk, B. and Maharaj, M. *Cyber Conflict: Competing National Perspectives*. John Wiley & Sons, 2013. ISBN 1848213506.

VanderSloot, B., Amann, J., Bernhard, M., Durumeric, Z., Bailey, M., and Halderman, J. A. *Towards a Complete View of the Certificate Ecosystem*. IMC, 2016.

Vegh, S. *Hacktivists or Cyberterrorists? The Changing Media Dscourse on Hacking*. *First Monday*, 7(10), 2002.
URL `http://ojs-prod-lib.cc.uic.edu/ojs/index.php/fm/article/view/998/919`

Venter, L. M., Wangwe, C. K., and Eloff, M. *e-Government Readiness: An Information Security Perspective From East Africa*. IST-Africa 2009 Conference Proceedings, 2009.

Vere, A. *Legal and Regulatory Frameworks for the Knowledge Economy. UN Documents, E/ECA/CODIST*, page 7, 2009.

Webb, G. *Evolution of Cyber Attacks Infographic*. 2013. Date Accessed 03 December 2015.
URL `https://www.venafi.com/blog/post/evolution-of-cyber-attacks-infographic/`

Wichers, D. *Open source web application security project top-10 2013. OWASP Foundation*, February 2013.

William, J. and Wichers, D. *Top Ten 2013 :The Ten Most Critical Web Application Security Risks*. 2013.
URL `https://www.owasp.org/index.php/Top_10_2013\discretionary{-}{}{}Top_10`

Winn, J. K. *Governance of Global Mobile Money Networks: The Role of Technical Standards. Washington Journal of Law, Technology & Arts*, 8(3):197–244, 2013.

Winterfeld, S. and Andress, J. *The Basics of Cyber Warfare: Understanding the Fundamentals of Cyber Warfare in Theory and Practice*. Syngress Publishing, 2012.

Workman, M. *Gaining Access with Social Engineering: An Empirical Study of the Threat. Information Systems Security*, 16(6):315–331, 2007.

Wu, M., Miller, R. C., and Garfinkel, S. L. *Do Security Toolbars Actually Prevent Phishing Attacks?* In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 601–610. ACM, 2006.

Wustrow, E., Karir, M., Bailey, M., Jahanian, F., and Huston, G. *Internet Background Radiation Revisited.* In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, pages 62–74. ACM, 2010.

Xia, H. and Brustoloni, J. C. *Hardening Web Browsers Against Man-in-The-Middle and Eavesdropping Attacks.* In *Proceedings of the 14th international conference on World Wide Web*, pages 489–498. ACM, 2005.

Yates, D. *A System for Characterising Internet Background Radiation.* Honours thesis, Rhodes University, 2014.
URL `http://www.cs.ru.ac.za/research/g11y1408/thesis.pdf`

Yazar, Z. *A qualitative risk analysis and management tool. SANS InfoSec Reading Room White Paper*, 2002.

Yilek, S., Rescorla, E., Shacham, H., Enright, B., and Savage, S. *When Private Keys are Public: Results from the 2008 Debian OpenSSL Vulnerability.* In *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*, pages 15–27. ACM, 2009.

Zandbelt, J., Hulsebosch, R., Bargh, M., and Arends, R. *Trusted Directory Services for Secure Internet Connectivity: Transport Layer Security using {DNSSEC}* . *Electronic Notes in Theoretical Computer Science*, 197(2):91 – 103, 2008. ISSN 1571-0661. Proceedings of the 3rd International Workshop on Security and Trust Management (STM 2007).

Zitser, M., Lippmann, R., and Leek, T. *Testing Static Analysis Tools Using Exploitable Buffer Overflows from Open Source Code.* In *ACM SIGSOFT Software Engineering Notes*, volume 29, pages 97–106. ACM, 2004.

Appendix A

Code and Scripts

This Appendix contains listings containing lines of code and scripts that were used to plot the graphs in Chapters 5 and 6 but also for data filtering whose data has been used throughout this research.

A.1 Tcpdump Packet Filtering Scripts

This is the script that was used to filter the data from three of Rhodes University's network telescopes. The three distinct top-level IP version 4 network address blocks are : 146/8, 155/8 and 196/8. Apart from the IP addresses in the script and the paths that hold specif packets from a specific network telescope the script operational principal is the same for all four countries thus the one below is for demonstration purposes. Kenya script was selected because it has a bigger IP address space allocated to it.

Following this script which was another script called which was used to executed the packet filtering script. It is a case of a script being executed by another script where the new file produced is a filtered version of the original script (containing only Kenyan data). The new files formed were smaller compared to the original file as it only accommodated traffic in East African coming from Kenya, Tanzania, Malawi and Uganda

A.2 tshark commands

Network telescope collects data in .pcap format as such there was need to convert the data into .csv with the the purpose of efficiency and ease of data manipulating with .csv

Listing A.1: Tcpdump data filtering script

```
!/bin/bash
DATA='/mnt/datapool/telescope/process/196.X.X'
OUT='/home/telescope/kenya'
/usr/sbin/tcpdump -nr $DATA/$1 net 5.10.82.0/29
or net 5.10.84.0/32 or net 5.153.49.0/29 or net 37.58.121.0/28
or net 41.57.96.0/20 or net 41.72.160.0/19 or net 41.75.32.0/20
or net 41.75.144.0/20 or net 41.75.176.0/20 or net 41.76.168.0/21
or net 41.76.184.0/21 or net 41.78.24.0/22 or net 41.78.176.0/22
or net 41.79.8.0/22 or net 41.79.168.0/22 or net 41.79.228.0/22
or net 41.79.252.0/22 or net 41.80.0.0/15 or net 41.89.0.0/16
or net 41.90.0.0/16 or net 41.138.240.0/20 or net 41.139.128.0/17
or net 41.190.252.0/22 or net 41.191.192.0/21 or net 41.203.208.0/20
or net 41.204.160.0/19 or net 41.206.32.0/19 or net 41.207.64.0/19
or net 41.207.96.0/19 or net 41.209.0.0/18 or net 41.212.0.0/17
or net 41.215.0.0/17 or net 41.215.128.0/20 or net 41.215.192.0/20
or net 41.217.220.0/22 or net 41.220.112.0/20 or net 41.222.8.0/21
or net 62.128.160.0/20 or net 77.220.0.0/19 or net 80.72.96.0/20
or net 80.231.19.0/26 or net 80.231.210.0/24 or net 81.199.4.0/24
or net 81.199.115.0/27 or net 82.206.136.0/24 or net 82.206.143.0/29
or net 82.206.239.0/24 or net 87.255.96.0/19 or net 104.166.101.0/24
or net 104.222.220.0/24 or net 105.48.0.0/12 or net 105.160.0.0/13
or net 105.230.0.0/15 or net 154.70.0.0/18 or net 154.72.0.0/22
or net 154.72.28.0/22 or net 154.76.0.0/14 or net 154.118.232.0/21
or net 154.122.0.0/15 or net 155.254.254.0/24 or net 159.8.96.0/30
or net 159.8.98.0/28 or net 159.8.163.0/28 or net 159.8.217.0/28
or net 159.8.224.0/28 or net 165.90.0.0/19 or net 169.239.160.0/22
or net 169.239.168.0/22 or net 169.239.172.0/22 or net 196.6.202.0/23
or net 169.255.8.0/22 or net 169.255.96.0/22 or net 169.255.104.0/22
or net 169.255.200.0/22 or net 169.255.212.0/22 or net 193.109.66.0/23
or net 194.9.64.0/23 or net 194.9.82.0/23 or net 195.202.64.0/19
or net 196.1.132.0/24 or net 196.3.58.0/23 or net 169.239.252.0/22
or net 196.6.215.0/24 or net 196.6.218.0/23 or net 196.6.220.0/24
or net 196.11.88.0/23 or net 196.11.190.0/23 or net 196.13.121.0/24
or net 196.13.136.0/23 or net 196.13.173.0/24 or net 196.13.191.0/24
or net 196.13.209.0/24 or net 196.13.255.0/24 or net 196.22.131.0/24
or net 196.28.11.0/24 or net 196.32.226.0/23 or net 196.40.128.0/20
or net 196.41.68.0/24 or net 196.43.192.0/24 or net 196.43.202.0/24
or net 196.43.205.0/24 or net 196.43.211.0/24 or net 196.43.212.0/24
or net 196.43.217.0/24 or net 196.43.218.0/24 or net 196.43.220.0/24
or net 196.43.222.0/24 or net 196.43.228.0/24 or net 196.43.239.0/24
or net 196.43.245.0/24 or net 196.43.246.0/24 or net 196.43.248.0/24
or net 196.46.16.0/24 or net 196.96.0.0/12 or net 196.200.16.0/20
or net 196.200.32.0/20 or net 196.201.128.0/19 or
or net 196.201.224.0/22 or net 196.202.160.0/19
or net 217.199.144.0/20 -w $OUT/$2
```

Listing A.2: Script that execute tcpdump script

```
#!/bin/bash
./kenya.sh 2009/2009.pcap kenya2009.pcap
./kenya.sh 2010/2010.pcap kenya2010.pcap
./kenya.sh 2011/2011.pcap kenya2011.pcap
./kenya.sh 2012/2012.pcap kenya2012.pcap
./kenya.sh 2013/2013.pcap kenya2013.pcap
./kenya.sh 2014/2014.pcap kenya2014.pcap
./kenya.sh 2015/2015.pcap kenya2015.pcap
```

Listing A.3: Data conversion tshark commands

```

#!/bin/bash
./tsharkScrip.sh kenya2009.pcap kenya2009.csv
./tsharkScrip.sh kenya2010.pcap kenya2010.csv
./tsharkScrip.sh kenya2011.pcap kenya2011.csv
./tsharkScrip.sh kenya2012.pcap kenya2012.csv
./tsharkScrip.sh kenya2013.pcap kenya2013.csv
./tsharkScrip.sh kenya2014.pcap kenya2014.csv
./tsharkScrip.sh kenya2015.pcap kenya2015.csv
```

than pcap. The following commands were used to do the conversion. As it ha been shown with the previous codes , Kenya data has been used as a demonstrating case once more.

A.3 Source code for graph Plots

To come up with graphs used in the subsequent Chapters 5 and 6 python was used, with comments attached to it to explain the rationale behind every line of code used. For the arrays, the values were changed to suit the country of interest. In addition to this, the port numbers were also customised depending on what was registered for each country but also the time period in which data was recored. The first graph accommodated the first four years and the second graph accommodates the remaining three years (2013,2014 and 2015). This was done for presentation purposes since the graph could not accommodate all the seven years.

The legend is colour coded in a way that each colour represents a separate year. This is to say that the bar graphs are grouped by first port number, then by year. This was chosen in order to show progress of each port as to show differences that have happened

over time with the ports of interest. It is worth noting that the code shown in Listing 4 accommodates all the seven years but during plotting it was split into two i.e two merged codes are shown in the code covering the seven year period and not split into two as the graphs will later show.

The only things that changed each time the code was run are the labels used on the legend, the port number and the values in the arrays all of which was done to accommodate the differences and the changes in different period and countries.

A.4 AWK Used for Shodan Formating

To format the and eliminate the anomalies found the SHODAN data sets, square brackets ("[" and "]") were added at beginning and end followed by commas between each entry. Vim commands gg goes to the first line, "[" and "]" selects the whole line G #Go to the last line where :s/$/'/ was applied reg-ex to selection. This removed the irregularities that came along with the data and made readable by python and and AWK

A.5 Python Code for Shodan Data Extraction

The code in this section was used for data extraction from the original SHODAN file after it was formatted with AWK command. Note that the characteristics of interest in the code are not arranged in any specific order primarily because of fitting purposed otherwise it is meant to follow the same pattern from first line to last one. This does not mean that the results came out differently i.e. the variables were still the same. The new file formed is the one that was used to carry out analysis while the records printed on the screen were used to verify progress of the script as the files were over 100MB each making it run slow, however this offered immediate feedback of what the script was doing.

Listing A.4: Python graph plotting

```
import numpy as np
import matplotlib.pyplot as plt

#declaring arrays with thier respective values
Year_2009 = [0,4,74,3673,2]
Year_2010 = [792,0,111,852,48]
Year_2011 = [509,0,51,3141,0]
Year_2012 = [1754,524,0,5088,484]
Year_2013 = [510,960,1177,1566,3479,255]
Year_2014 = [7547,2301,623,3437,3367,1067]
Year_2015 = [2776,680,0,160,1449,339]

#declaring the number of arguments accepted by the arrays
n_groups = 5 ind = np.arange(n_groups)

# create plot
fig, ax = plt.subplots()
bar_width = 0.24
opacity = 0.8

#create atleast three bargraphs on one one plot
#colour coding the graphs, legend and adjusting the graphs width

rects1 = plt.bar(ind,Year_2009, bar_width,alpha=opacity,
                 color='b',label='2009')

rects2 = plt.bar(ind + bar_width,Year_2010, bar_width,
              alpha=opacity,color='r',label='2010')

rects3 = plt.bar(ind + bar_width +0.25,Year_2011,bar_width,
                 alpha=opacity,color='g',label='2011')

rects4 = plt.bar(ind + bar_width +0.50,Year_2012,bar_width,
            alpha=opacity,color='k',label='2012')

#labeling of the graph axes and title
plt.xlabel('Port_number')
plt.ylabel('Port_count')
plt.title('Port_number vs Port_Count')
plt.xticks(ind + 0.4, ('22','23','1433','445','3389'))

#display the bar-graph and legend on the plot
plt.legend()
plt.tight_layout()
plt.show()
```

Listing A.5: Python data extraction code

```
import json
#read from file using the path given
f = open('/home/c7469/Documents/masters/olivier/
shodan_data_CVE.json','r')
g = json.load(f)

#read records and print those that fit the criteria on the screen
for i in range(len(g)):
    if g[i]['isp'] != "" and g[i]['port'] != "" and
       g[i]['os']!= ""and  g[i]['_shodan']['product'] != ""and
g[i]['devicetype']!="" and g[i]['Openssl']['version'] != "":
    print(g[i]['isp'], g[i]['port'], g[i]['_shodan']['product'],
        g[i]['devicetype'],g[i]['os'],g[i]['Openssl']['version'])

#read records and write those that fit the criteria to a new file
with open("/home/c7469/Documents/masters/olivier/
shodan_data_CVE_edited.json", "w") as outfile:
     for i in range(len(g)):
         if g[i]['isp'] != "" and g[i]['port'] != "" and
            g[i]['os']!= "" and  g[i]['_shodan']['product'] != ""
            and g[i]['devicetype']!="" and
            g[i]['Openssl']['version'] != "":
            outfile.write(json.dumps((g[i]['isp'],g[i]['port'],
            g[i]['devicetype'],g[i]['_shodan']['product'],
            g[i]['os'], g[i]['Openssl']['version'] ),outfile,
            indent=4,ensure_ascii=False))
```

Appendix B

Data

This Appendix contains a list of tables that contain additional data which was used for the analysis of this research. The data includes data sets A, B, C, D, E, F, G and H extracted from June 2015 to November 2016 SHODAN data and IBR data collected from January 2009 to February 2015.

Table B.1: Device types from SHODAN data set A

	Tanzania	Malawi
Device Name	**No. of Devices**	**No. of Devices**
Firewall	62	10
Media Device	2	1
PBX	35	-
Printer	111	23
Printer	6	-
Server Router	9	1
Switch	414	45
WAP	5	-
Total	644	80

Table B.2: ISPs in Tanzania from SHODAN Data 2015

ISP	**No. of Instances**
Aptus Solutions Ltd	863
Cats-Net Limited	372
Habari Node Ltd	732
Simba net (T) Limited	1432
Spice Net Tanzania Ltd	1230
Startel (T) Ltd	1070
TTCL	3376
University of Dar-Es-Salam	530
WIA Tanzania	729
Zanzibar Telecom(Zantel)	733
Other ISPs	4446
Total	15513

Table B.3: ISPs in Tanzania from SHODAN Data 2015

ISP	**No. of Instances**
Airtel	108
Access	30
Burco	39
Broadband	284
Globe-as	562
Malswitch	11
MAREN	77
MTL	976
Sky-band	1732
TNM	4
NIC.mw	13
Total	3836

Table B.4: Port numbers from SHODAN Data 2015

			Tanzania	Malawi
Port Number	**Protocol**	**IANA Names**	**Port count**	**Port count**
7	TCP	Echo	233	94
21	TCP	FTP	338	87
23	TCP	Telnet	2051	311
25	UDP	SMTP	272	57
53	UDP	DNS	869	349
80	TCP	HTTP	3675	663
110	TCP	POP3	233	94
111	TCP	SUN RPC/NFS	128	27
123	TCP	NTP	248	20
137	TCP	NETBIOS	267	59
143	UDP	IMAP4	196	57
161	TCP	SNMP	690	-
166	UDP/TCP	Sirius System	-	99
443	TCP/UDP	HTTP over SSL	1058	1279
500	TCP/UDP	ISAKMP	742	133
993	TCP	IMAP over SSL	76	13
995	TCP	POP3 over SSL	69	11
1723	TCP	MS PPTP	190	55
1900	TCP/UDP	SSDP	134	8
3389	TCP	MS-RDP	287	97
4500	TCP/UDP	IPSec NAT-Travel	705	119
7547	TCP/UDP	DSL Forum CWMP	1140	-
8443	TCP/UDP	PCSync HTTPS	69	-
8080	TCP	HTTP Alternate	954	85
9001	TCP/UDP	Other	229	-
9002	TCP/UDP	Other	67	-

Table B.5: Port numbers from SHODAN Data 2016

		Kenya	Tanzania	Uganda	Malawi
Port Number	**Protocol**	**Port count**	**Port count**	**Port count**	**Port count**
443	TCP/UDP	4492	1350	1055	986
23	TCP	5760	2246	1029	615
21	TCP	2374	1034	170	457
25	UDP	1160	317	152	74
110	TCP	676	217	113	94
80	TCP	7112	3356	1789	753
8080	TCP	2094	800	258	69
22	TCP	3852	2096	1494	664
4500	TCP	1168	506	228	195
500	TCP/UDP	1717	752	402	204
161	TCP	1678	559	2441	151
2000	TCP	2276	2270	358	924
3389	TCP	714	110	45	90
Total		**43206**	**11119**	**13244**	**4563**

Table B.6: Kenya's IBR data

Time line of data	D-port	No. of Ports	IP Source	No. of SRC_IP
2009	445	16412	196.201.141.x	8142
	1433	297	196.201.148.x	4570
	135	1118	196.200.21.x	1048
	25	11385	196.201.208.x	980
	23	463	196.202.195.x	598
Total count		**29945**		**29945**
2010	21	503	196.201.208.x	2330
	22	1159	196.202.217.x	1800
	135	7353	196.202.212.x	1478
	445	13206	196.201.226.x	1004
	1433	561	62.24.111.x	821
Total count		**24848**		**23462**
2011	135	46	41.139.193.x	2077
	139	1016	196.202.196.x	972
	1433	576	41.204.167.x	563
	22	2767	41.89.160.x	595
	445	5137	212.49.95.x	512
Total count		**9894**		**9894**
2012	22	751	41.139.255.x	1531
	23	577	196.202.214.x	1128
	445	4133	41.204.168.x	1079
	5900	718	41.204.186.x	768
	3389	1171	217.199.145.x	957
Total count		**14441**		**14441**
2013	22	956	197.232.13.x	2468
	80	2056	196.201.211.x	3072
	445	7036	196.201.208.x	1536
	3389	719	196.200.16.x	524
	8080	1280	196.202.202.x	599
Total count		**15815**		**15815**
2014	22	5699	212.49.70.x	6140
	23	20252	41.215.28.x	1540
	25	5151	41.222.15.x	10178
	445	4171	41.57.99.x	851
	3389	3275	41.57.105.x	811
	5900	1024	197.211.12.x	995
Total count		**41776**		**41776**
2015	22	3307	197.248.144.x	2288
	5916	913	41.203.214.x	2823
	80	2858	197.232.26.x	1136
	445	2444	212.49.70.x	2913
	8080	2077	197.248.168.x	852
Total count		**15133**		**15133**

Table B.7: Tanzania's IBR data

Time line of data	D-port	No. of Ports	IP Source	No. of SRC_IP
2009	1433	74	196.44.161.x	896
	445	3673	196.45.39.x	291
	5207	12	196.43.78.x	246
	23	4	196.46.129.x	239
	37366	3	196.44.171.x	209
Total count		**3795**		**3795**
2010	139	49	196.44.161.x	539
	1433	111	196.43.67.x	518
	22	792	196.45.146.x	335
	3389	48	41.223.4.x	991
	445	852	41.221.41.x	255
Total count		**4244**		**4244**
2011	445	3141	41.220.180.x	499
	443	256	196.46.120.x	400
	22	509	196.43.84.x	328
	1433	51	196.44.162.x	256
	25	25	196.1.53.x	252
Total count		**4304**		**4304**
2012	445	5088	41.77.228.x	744
	3389	484	41.59.13.x	514
	23	524	196.46.122.x	510
	22	1754	196.45.156.x	741
	47886	11	196.44.173.x	512
Total count		**8071**		**8071**
2013	445	3479	196.45.144.x	2257
	80	960	41.216.216.x	768
	135	1177	41.59.7.x	1286
	22	510	196.43.84.x	1187
	3389	1566	196.41.43.x	700
Total count		**9723**		9723
2014	8080	1022	41.93.45.x	3072
	80	2977	196.41.43.x	3153
	445	3367	196.46.100.x	7781
	3389	3437	41.59.61.x	861
	22	5591	41.77.228.x	440
Total count		**19812**		**19812**
2015	445	1449	196.41.50.x	1353
	22	2776	196.46.100.x	2694
	8080	339	41.77.228.x	215
	80	680	196.45.144.x	364
	23	258	196.41.47.x	304
Total Count		**5897**		**5897**

Table B.8: Uganda's IBR data

Time line of data	D-port	No. of Ports	IP Source	No. of SRC_IP
2009	445	6819	196.0.13.x	2678
	135	5925	196.0.11.x	2361
	80	175	196.0.18.x	752
	22	47	196.0.5.x	881
	12692	3	196.0.17.x	3641
Total count		**13002**		**13002**
2010	445	14041	196.0.26.x	13422
	135	4258	196.0.19.x	2063
	22	24	196.0.12.x	760
	443	12	196.0.25.x	753
	25	758	196.0.13.x	308
Total count		**19215**		**19215**
2011	445	2459	196.0.4.x	769
	80	84	196.0.5.x	311
	1433	152	81.199.21.x	152
	19216	3	196.0.0.x	147
	13192	3	196.43.133.x	147
Total count		**2704**		**2704**
2012	445	3811	196.0.4.x	1711
	3389	48	196.0.41.x	1574
	22	920	196.0.26.x	341
	23	256	196.0.7.x	531
	25	2	41.210.184.x	309
Total count		**5060**		**5060**
2013	3389	142	196.0.4.x	270
	445	1215	41.223.85.x	220
	110	14	196.0.31.x	231
	25	231	196.0.64.x	171
	5900	75	41.221.84.x	153
Total count		**1721**		**1721**
2014	3303	512	212.88.100.x	418
	7778	239	196.0.29.x	768
	445	518	196.0.64.x	181
	3389	227	196.0.26.x	239
	23	11	41.84.196.x	148
Total count		**2218**		**2218**
2015	22	765	154.0.130 .X	768
	23	99	41.221.84.x	101
	3389	150	196.0.64	73
	445	195	212.88.112.x	44
			154.0.134.x	22
Total Count		**1204**		**1204**

Table B.9: Malawi's IBR data

Time line of data	D-port	No. of Ports	IP Source	No. of SRC_IP
2009	12203	295	41.221.106.x	1249
	2967	292	41.221.96.x	14
	2968	287	196.45.190.x	43
	5900	287	-	-
	445	155	-	-
Total count		**1331**		**1331**
2010	12203	428	196.216.8.x	258
	5900	417	41.221.106.x	1668
	2968	412	41.221.99.x	210
	2967	402	41.221.97.x	38
	22	258	-	-
Total count		**2215**		**2215**
2011	445	16	196.216.10.x	12
	3389	4	-	-
	3306	1	-	-
Total count		**21**		**14**
2012	445	63	41.221.96.x	8
	3389	40	41.221.96.x	67
	22	16	41.77.12.x	24
	3306	2	41.78.57.x	10
	80	1	-	-
Total count		**146**		**146**
2013	445	250	41.221.97.x	210
	80	255	105.234.255 .X	14
	3389	52	196.216.15.x	10
	25	3	41.77.13.x	255
	-	-	41.87.10.x	53
Total count		**580**		**580**
2014	445	240	41.221.97.x	197
	3389	5	41.77.14.x	12
	-	-	41.221.108.x	5
Total count		**262**		**262**
2015	139	115	41.221.109.x	8
	445	35	41.221.103.x	150
Total Count		**151**		**151**