PROTECTING INSTITUTIONS OF HIGHER LEARNING IN KENYA: A SCALABLE HYBRID DECOY FRAMEWORK AGAINST CYBER THREATS

EDWIN KIPRONO SEREM

A RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF MASTER OF SCIENCE IN COMPUTER SCIENCE OF THE UNIVERSITY OF EMBU

SEPTEMBER, 2021

DECLARATION

This research project is my original work and has not been presented elsewhere for a degree or any other award.

Signature…… Date……

Edwin Kiprono Serem

Department of Mathematics, Computing and Information Technology

B529/1158/2017

This research project has been submitted for examination with our approval as the University Supervisors

Signature…… Date……

Dr. David Muchangi Mugo

Department of Mathematics, Computing and Information Technology

University of Embu.

Signature…… Date……

Dr. Boaz Kipyego Too

Department of Mathematics, Computing and Information Technology

University of Embu.

DEDICATION

I sincerely want to dedicate this research work to my family members, my wife Carina, mother Jane, and siblings for the support and prayers during this period. May the Almighty God bless them all.

ACKNOWLEDGEMENT

I want to thank the Department of Mathematics, Computing and information technology, University of Embu, for mentorship and support. In addition, I would like to thank my supervisors, Dr. David Mugo and Dr. Boaz Too, for guidance and direction in the research study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABBREVIATIONS AND ACRONYMS

ACyDS	Adaptive cyber deception system
AIDS	Anomaly-based Intrusion Detection System
APT	Advanced Persistent Threat
CSP	Communications Service Provider
DDoS	Distributed Denial of Service
DDT	Decoy Distributor Tool
HIPS	Host-based intrusion prevention system
ICT	Information Communication Technology
IDS	Intrusion Detection System
LXC	Linux Containers
SIDS	Signature-based Intrusion Detection System
SIEM	Security Information and Event Management
SDN	Software-Defined Networking
VM	Virtual Machine

ABSTRACT

Cybersecurity threats are malicious acts that seek to damage, steal, or gain unauthorized access to information. Higher institutions of learning in Kenya have adopted the use of information systems in their service delivery. However, their level of preparedness to deal with emerging threats in their cyberspace is limited by techniques used to detect, inform, and deflect the cyber threats before they cause much harm. The main objective of this research study was to develop a scalable decoy framework for use in institutions of higher learning. The research process was done in two phases; the first phase encompassed preliminary studies that involved soliciting responses from 84 ICT personnel drawn from 42 institutions in Kenya selected through the purposive sampling method. This study made use of primary data collected using structured questionnaires, then descriptively analyzed. The findings revealed the institutions recorded cyber attacks within twelve months of the research period, and the main tools and techniques in place are inefficient to detect significant threats. The second phase entailed designing the framework prototype using Linux containers as decoys in the front and back end and monitoring the attacks using HonSSH, while graphical presentation used Grafana. The decoys were set in a layered approach. The front-end decoy hid the back-end decoy by internally configuring the front-end decoy to capture and reroute the attacker commands via a secure tunnel. The back-end decoy did the processing of commands issued through the front-end decoy then gave feedback. Simulation of user activities and network traffic generation was achieved using the General HOSTS framework to make it more realistic to the attacker. The attacker's virtual machine used Kali Linux. Scalability, latency, and throughput metrics were used to test the framework's effectiveness; decoy data analysis was done by logstash and pipelined to Kibana for visualization. The experimental results demonstrate that the system effectively misdirected commands by combining deceptive network setup and configurations and generating fake user and network activities with an average latency of 0.0015s, throughput 864Mbits/s, and boot speed 7.485s. The study highly recommends including cyber decoys in the institutions network to boost security in a proactive approach due to effectiveness in utilizing computing resources. The framework will help cybersecurity professionals protect higher institutions of learning from stealthy and sophisticated attacks. This research work contributes to knowledge in designing and developing effective deceptive decoys tools in cybersecurity research.

CHAPTER ONE

INTRODUCTION

1.1. Background information

Cyber threats are computer-based actions or systems accused of harming other systems and networks by damaging data, breaching authorization protocols, and denying access to information. Cyber threats include computer malware, breach of data, and service interruptions (Heckman et al., 2015). It can also be defined as the success of a cyberattack to gain access to unauthorized information, interrupt systems, damage, and steal information assets. It can also extend to intellectual property or other sensitive data (Klenka, 2021). These threats originate from trusted users within an organization or through unknown parties over remote access.

The sophisticated cyber threats significantly advanced persistent threats (APT) are the most lethal in the 21st Century (Johansson, 2019). These attacks are well-planned and launched by experienced attackers with good knowledge of the existing cyberinfrastructure such as firewalls, Intrusion Detection Systems (IDS), and Intrusion Prevention Systems (IPS). Their success is attributed to exploiting well-known vulnerabilities and conducting social engineering-related campaigns.

Cybersecurity is the defence mechanism in protecting servers, personal computers, corporate networks, data, and electronic systems from malicious attacks. This practice involves both information security and electronic assets security against destructive cyber-attacks (Kaspersky, 2019). Defences in several enterprise networks consist of static defences such as antivirus endpoint firewalls or anti-malware, which effectively prevent known attacks.

Due to the attacks' lethal nature, it is vital to implement security defences that are proactive and adaptive to ever-evolving attacks (Cranford et al., 2020). One such form of protection is proactive deceptive decoys systems deployed together with natural systems in the network. Deception decoys are some of the new security defence mechanisms under research. If well utilized, they can enhance security by diverting the

cyber attackers from the natural systems and improving the practical study of the attack tools under different setups. Deception is the key ingredient that makes the attacker's job difficult due to the complex design of the decoys.

The cyber defender’s job is difficult, but defensive cyber deception is a promising research area that might bring some advantage back to the defender (Heckman et al., 2015). Cyber deception techniques are vital in slowing down attacks using misinformation, delays, and determent tools, making the deception techniques more effective and efficient in cyber defence. These techniques include decoy systems, honeypots, and tar pitting (Fugate & Ferguson-Walter, 2019).

1.1.1. Honeypots

Honeypot is a common deception technique; it mainly lures attackers to false networks or systems. They are also used to contain and monitor attackers and their related activities (Provos, 2004). The honeypot techniques have been enhanced by creating a false network design and presenting it as an actual topology (Trassare et al., 2013). Cohen and Koike (2003) made a similar deception framework to analyze a wide range of deceptions involving people and computers. Aggarwal et al.(2016) tested the effectiveness of honeypots using cybersecurity games. Their research revealed how attackers manoeuvre through different setups. Due to the evident successful implementation of honeypots, reals systems have been made look fake using fake honeypots (Rowe et al., 2007).

1.1.2. Decoy Systems

Decoy systems use deception techniques though they differ from honeypots regarding technology and function (Bringer et al., 2012). The unique characteristic of the decoy systems is that they are deployed within the existing network as part of the systems in the network. What makes them more robust is the agility that comes with their configurations that make typical networks look heterogeneous and low-fidelity (Ferguson-Walter et al., 2017). One advantage of the decoy is lower system maintenance compared to full honeypots.

Unlike conventional IDS, decoys have the advantage of exposing stealthy attacks effectively, considering that legitimate users usually are not required to log in to the decoys (Vasilomanolakis et al., 2015). Besides, decoys are known to accelerate information gathering on attacks (Beham et al., 2013) using dedicated traps, which misinforms attackers with falsified information such as fake passwords, documents, and encryption keys (Pfleeger & Stolfo, 2009). Decoy networks or systems may effectively convince attackers into believing that they have succeeded in penetrating the existing system or network by the data relayed by the decoy. In contrast, in the real sense, they have only penetrated one or more of the decoy mirage nodes in the complex framework.

1.2. The cyber security situation in Kenya

Cyber security is an emerging trend in the information technology industry in Kenya. It is attributed to the internet and automation in the economic sectors (Oyelaran-Oyeyinka & Adeya, 2004). Higher learning institutions are the leading research and academic centres; information and knowledge are shared online or over intranet systems. This form recipe for threat actors to either steal the information or deny access to information in the institutions depending on their intention. Critical information is stored in the systems, such as the student records, medical records, marks, and other confidential information. Other information found in the institutions is the staff details, research work, and confidential documents.

The Communication Authority, through the Kenya Computer Incident Response Team – Coordination Centre (National KE-CIRT/CC), is a multi-agency collaboration framework that is responsible for the national coordination of cyber security as Kenya's national point of contact on cyber security matters (Fielder, 2021) reported that there has been a sharp increase in cyber threats during the period October-December 2020, with 56,206,097 incidents detected, This was a 59.8% increase from the previous period July to September 2020, where 35,173,937 cyber threat events had been detected. There was a significant increase in malware attacks at 46 million, followed by web application attacks at 7.8 million. At the same time, 2.2 million Distributed Denial of Service (DDoS) threats were detected during the same period.

Institutions of Higher Learning form part of this cyberspace due to the use of the internet in research, storage, and use of related system services (Maranga & Nelson, 2019). These results in advanced cyber threats, especially in financially related data and examination data. In Kenya, for instance, the emergence and successful use of mobile banking and Mpesa services catapult cyber cases and fraud that target education institutions. It is attributed to system-bank integration and weak system configuration to withstand the attacks (Chetalam, 2018).

There are at least four reasons institutions of higher learning are the target for cybercriminals (Chapman, 2019). These include data theft, where cyber criminals get valuable information then sell them or, in other circumstances, use the information as a blackmail tool on their victims. The second motive is a financial benefit to intercept key transactions between the institution's systems and the banks. The third reason is espionage. The institutions of higher learning such as universities and colleges are centres for research and technical inventions, hence the target for cybercriminals due to the intellectual property, which is valuable and expensive projects. This attack was launched to steal intellectual property. This study focused on addressing the shortcomings of the existing cyber systems in the institutions of higher learning by implementing an adaptive decoy system.

1.3. Statement of problem.

Several institutions of higher learning in Kenya have adopted the use of information technology with robust systems and online services without proper cybersecurity parameters (Serianu, 2020). Most rely on static cyber defences while cyber attacks keep evolving day by day. The existence of static cyber defence infrastructure cannot deter attackers from accessing the protected network due to the emergence of advanced penetration tools and techniques that occasionally has rendered the static cyber defence ineffective and inefficient (Jingyao et al., 2020).

Another problem is detecting attacks in real-time and analyzing them for better service delivery in institutions of higher learning (Ferguson-Walter et al., 2019) noted that most commercially available decoys systems are not adaptive, usually preconfigured static cyber defence tools. These setups in the pre-existing systems and research on adaptive

defence show that they can bear good fruits if properly designed and developed in the future. Chiang et al. (2016) came up with an adaptive cyber deception system (ACyDS) which provided a unique virtual network view to each host in an enterprise network.

Based on the state of the art literature, it was noted that existing systems lack believability due to inadequate user activities and network traffic in the decoys, which makes them easy to identify from the existing systems. In addition, they lack proper log management of the events in an active cyber security incident, which is vital in analysis. This research came up with an adaptive deception decoy system using a layered approach while introducing simulation of user activities and network traffic in the decoys and reporting to alert cyber defenders of the system's status in real-time.

1.4. Research questions

i. What are the existing cyber threats analysis and evasion techniques used in institutions of higher learning in Kenya?
ii. How will simulation of the user and network activities be achieved in an adaptive hybrid decoy system?
iii. How effective and efficient is the hybrid adaptive decoy system framework?

1.5. Objectives

1.5.1. General objective

The general objective of the study is to develop an adaptive defensive cyber decoy framework.

1.5.2. Specific objectives

i. To determine the existing cyber threats analysis and evasion techniques used in institutions of higher learning in Kenya.
ii. To develop an adaptive hybrid deception decoy framework for network-based systems.
iii. To measure the effectiveness and efficiency of the adaptive hybrid decoy system framework.

CHAPTER TWO

LITERATURE REVIEW

2.1. Cyber threats detection and evasion techniques

Cyber threats detection is vital because it provides ways of evading cyber-attacks. Cyber-detection involves identifying the attacks before they occur and taking necessary security measures (de Bruijn & Janssen, 2017). On the other hand, cyber-evasion techniques refer to how systems evade attacks launched against the services. Cyber-attack detection has served as the primary method cybersecurity professionals use to mitigate the ever-evolving world of Advanced Persistent Threats (APTs) (Friedberg et al., 2015). Some of the critical techniques used are; Intrusion Prevention systems, Security Information and Event Management solutions, Intrusion Detection Systems, Antimalware, Firewalls, and the use of network policy control solutions (Thapa & Mailewa, 2020).

2.2. Intrusion Detection Systems

Intrusion in computer systems can be defined as unauthorized activity that causes harm to an electronic system when executed. These activities can deny authorized users access to the existing systems or unauthorized users execute harmful actions that harm the organization (Khraisat et al., 2019). An intrusion detection system is a hardware or software that detects malicious actions or processes in the systems which may not be recognizable by the existing tools like firewalls. Intrusion Detection Systems are categorized into Signature-based and Anomaly-based (Khraisat et al., 2019).

2.2.1. Signature-based Intrusion Detection System

Signature intrusion detection systems work by using already known information about attacks. When an attack occurs, the SIDS compares the patterns the attack uses, and the knowledge base then takes necessary action (Khraisat et al., 2019). Each intrusion has a signature stored in the database; therefore, an alarm is triggered when the signature matches the known intrusion. The ever-rising cases of zero-day cyber-attacks have

rendered techniques used in most SIDS ineffective to combat them due to a lack of existing signature on the attacks (Symantec, 2017).

2.2.2. Anomaly-based Intrusion Detection System.

An anomaly-based Intrusion Detection System was developed to address the challenges of the signature-based IDS. It is done using machine learning combined with knowledge-based methods or statistical techniques (Khraisat et al., 2019). Developments involve two key phases; the model training and the testing phase. The training phase involves regular traffic, which is profiled to learn a standard behaviour model; in training, a new data set is used to establish the system's capability to check on unforeseen intrusions. The advantage is identifying the zero-day cyber attacks since it is based on abnormal user activity rather than a signature database (Alazab et al., 2012).

An anomaly-based Intrusion Detection System has some benefits. The first is the ability to discover internal malicious activities. It does this by checking the transactions made by the intruder using stolen credentials that are uncommon in normal user activities; then, it creates an alarm. Second, it is tough for an intruder to know what a typical user behaviour looks like without raising an alert as the underlying system is custom profiled.

2.3. Security Information and Event Management System

Security information and event management system combines security information and event management systems, while the former deals with compliance with standards and policy by consolidating logs and data analysis. The latter focuses on technical support in real-time threat management, events, and security incidents (Tyagi, 2017).

Security information and event management system tools use event data gathering and data logged from the host's systems and applications. The systems then centralize all data into one platform. Data can originate from other security devices. SIEM tools sort data and categorize them as successful or failed logins(González-Granadillo et al., 2021).

Despite having various advantages, SIEM has a limitation in its operation. These are some of the limitations (González-Granadillo et al., 2021). First, implementation takes longer due to the technical support required to ensure security controls integrate with systems already in place in the organization. Secondly, it is expensive and requires highly talented experts to manage. Also, they depend on rule-based data analysis, which is ineffective due to many logged data. Finally, if misconfigured, it may lead to inefficiency.

2.4. Firewalls

Firewalls use traffic filtering in different zones of the network. These zones are untrusted zone, trusted zone, and Demilitarized Zone (DMZ). Some firewalls are host-based, which serves a single machine (Tyagi, 2017).

Firewalls have several limitations, including: need to deploy them in all points where organizations network connects to the outside world; being a host software in operating systems makes vulnerable to the weakness of the host. Another issue is that they focus on external attacks, leaving insider attacks unattended and lacking relevant reports and analysis of blocked packets due to rule mismatch.

2.5. Intrusion Prevention System

IPS works by preventing intrusion through a set of deny rules. It looks like an inverted firewall because it uses well-known security problems as essential rules in preventing intrusion. It detects malware activities using existing binary signatures, logs, and behavioural patterns showing security policy violations (Tyagi, 2017). When the packets show up, the IPS checks the list of rules indicating it should drop the packet.

2.5.1. Host-based IPS

A host-based intrusion prevention system is a single computer-based system usually designed to protect the host system. This kind of system prevents malicious code from altering the host system by picking and comparing the resulting changes then prohibit or send an alert for permission (Conrad et al., 2012). However, these systems have limitations due to challenges related to configuration and implementation. Another

challenge is related to pop-ups that can make it hard for users to deal with the use of the system due to frequent requests for permission to execute something (Tyagi, 2017).

2.5.2. Network-based IPS

Network IPS is a scaled-up intrusion prevention mechanism that uses sensors and monitoring tools to capture and analyse network traffic. The malicious activities are sensed in real-time then appropriate action is taken. The sensors are deployed in designated network points for network surveillance while it is occurring, irrespective of the location of the attack target (Paquet, 2012).

2.6. Honeypot

A honeypot is a captivating system designed and programmed to work like natural systems that are likely targets of cybercriminals (Perkins & Howell, 2021). It works effectively by detecting or deflecting cyberattacks from affecting the real target.

2.7. Decoys

cyber decoys are fake systems that mimic natural systems in a network with the same attributes as an entire system in behaviour and performance only that legitimate users cannot access them (Almeshekah & Spafford, 2016).

2.8. An adaptive hybrid decoy system model of a network-based system

Decoy systems form the key deceptive platform fundamentally to counter sophisticated cyber-attacks. Several cybersecurity-related research activities or projects have been done in the design and implementation of decoy systems. Sun, Liu and Sun, (2019) proposed a hybrid decoy system for constructing high-fidelity decoy networks to defeat remote malicious reconnaissance in the pre-exploitation phase and insider threats in the post-exploitation phase. Their research separated decoys into lightweight LXCs that run on a front-end decoy server and the back-end-server running executions of the commands and return the output to the front-end decoys. However, the research did not focus on the simulation of user activities in achieving the adaptability of the framework to ever-persistent attacks, especially in the post-exploitation phase. In addition, they

failed to capture the logs into a separate database for further analysis, which is a good practice for any cyber defence system in mitigating threats. Another problem with their approach is using virtual box VMs to host servers that cannot be deployed in an actual network setup.

Aggarwal et al. (2016) came up with a decoy framework in the form of a game. In their proposal, a deception game can be used to evaluate the decision-making of a hacker in the presence of deception. The main limitation of this research is that the focus was on effectiveness to the amount of deception used and the timing of deception. Another challenge is that all their work is purely conceptual and theoretical, which lack implementation therefore not able to reflect any existing natural system or network scenario. The scope of their project was limited to the assumption of the behaviour of the attacker and the defender, which in real sense it doesn't happen as perceived.

In another work, A scalable, high fidelity decoy framework against sophisticated cyber attacks. et al. (2016) came up with an adaptive cyber deception system (ACyDS) which provided a unique virtual network view to each host in an enterprise network. It was to mimic a real-time network topology by coming up with subnets and virtual appliances that look like an enterprise network.

Ferguson-Walter et al. (2019) developed a framework using cyber games whereby each player has absolute individual perception and moves taken in the game. Besides, their design provided that a player can manipulate the decision of the other, thus influencing them to take sub-optimal actions. This work was limited by the perception that attackers and defenders are only guided by available moves, which is not an actual attack scenario. Also, it lacks practical application in a hybrid decoy system because discussion only tackles attacker and defender in a singular form.

Further research was done by Vrable et al. (2005), which resulted in the development of Potemkin, a virtual honey farm that could support hundreds of VMs.It formed the basis of the design and development of a hybrid version of the honey farms. However, due to their limitations and identifiability, the honeypots may not fit the design of high-fidelity decoy systems.

Also, decentralized honey farms have been developed, which use virtualization in deployment (Jiang and Wang, 2006). Jiang and Wang proposed Collapsar, a virtual machine-based architecture for network attack capture and detention in this research work. Collapser could use a dedicated network designed locally and could support a large number of highly interactive honeypots.

A hybrid decoy system architecture was developed by (Sun et al., 2017). Their work developed a decoy-enhanced defence framework that can proactively protect critical servers against targeted remote attacks through deception., the system followed a hybrid architecture that separates lightweight yet versatile front-end proxies from back-end high-fidelity decoy servers to achieve high fidelity and good scalability. However, the research didn't succeed in generating believable user activities in the decoys and generating network traffic when a stealthy attack can quickly identify the decoy from the real servers due to the immobility of the internal services.

Hutchins et al. (2011) recommended using The Cyber Kill Chain framework to enhance proactive cyber defence. The framework, which derives its model from military attacks, illustrates the steps a threat actor takes from the beginning of the attack through completion. Threat actors move consecutively through the following phases: Reconnaissance, Weaponization, Delivery, Exploitation, Installation, Command & Control, and Actions on Objectives. By understanding the actions threat actors take to compromise an organization, cyber defenders can focus on moving proactive defences towards the initial phases in The Cyber Kill Chain. It’s a model often criticized for focusing on perimeter security like firewalls and limited to malware prevention, a standard limitation amongst cyber tools.

Wang and Lu (2018) noted that the cyber defences that are in use in the industry are designed to respond only when cyber threats are at an advanced stage, the Delivery phase of the Cyber Kill Chain. This argument concurs with the findings of Mirilla, (2018). He establishes that among the challenges encountered by an organisation with malicious software is detecting and stopping malware in the delivery stage. It entirely depends on the signatures used, detection of malware behaviour, or heuristic analysis

of the machine's attack by the malware. It makes the framework ineffective in handling the advance persistent threats.

Some researchers are focusing on Artificial Intelligence (AI) and Machine Learning (ML) to address the challenges of perimeter cybersecurity solutions. It involves the use of the intelligent system to understand and protect the machine autonomously with human intervention. One key researcher in this field is Dr. Christopher Whyte, assistant professor at Virginia Commonwealth University. He described the ramifications of the current cyber arms race in AI and ML. According to Whyte (2020), if data processing is done using a machine learning algorithm. Programmers dictate machine learning performance in terms of data ingestion, processing, and output. The challenge with the current products by various cybersecurity tools vendors is that they sell their products, claiming that they use machine learning and artificial intelligence. Still, in a real sense, they are not and only use the simple machine learning analysis that is less efficient to tackle persistent threats.

Further to the above weakness, the reliance on AI and ML in addressing the cyber defence has other limitations because the exact mechanisms used in machine learning and artificial intelligence as a technique to defeat the attackers are utilising the attackers in coming up with advance cyber threats more lethal that the algorithms used to design the defence. The result is that current defensive capabilities will continue to diminish as threat actors become more capable through machine-led attack strategies (Whyte, 2020).

Another argument was brought forward by Zhang et al. (2015), who outlined one of the problems with AI and ML is the culture of the hackers. The hacker generously shares the tools and techniques they use to attack and keep improving them due to democratic communities that share knowledge without prejudice. On the other hand, the organization hardly shares their knowledge with others due to privatization and related copyrights. It makes them weak and gives hackers the upper hand in both skills and knowledge.

Dlamini et al. (2020) introduced honey files as another defensive deception tactic. Honeyfiles provide fake data with the characteristics of legitimate data to lure threat

actors. The threat actor activity initiates detection mechanisms. Successful honey files detect data exfiltration and alert on unauthorized access. However, the limitation of the honey files lies in the effectiveness to convince the attackers into believing them because, in most cases, they are identifiable.

2.9. Simulation of user activities and network traffic in decoy systems

Decoy systems' effectiveness is determined by the level at which it manages deception at different levels of operations. One tool that makes the decoys look like the existing system is the simulation of the user activities and network traffic that adaptively lure the attacker into believing that they have access to the system. Typically decoys are static pre-configured systems that are easily identifiable by the attacker through a series of tools.

Esther et al. (2020) proposed a decoy prototype by simulating the electronic health record system data using a decoy-based system named HoneyDetails. HoneyDetails serves fake data to the attacker in an active cyber attack. However, the adversary will be convinced by the nature of the data that looks more realistic. This prototype showed that data simulation in an electronic health record system could safeguard patients' data. However, the limitation was on decoy generation in the medical domain and the inability to defend the system proactively.

In another research work, Niels and Holz (2008) developed Honeyd, a virtual honeypot that simulated TCP/IP stack in an operating system. The only limitation was the lack of isolation of the stacks; thus, the entire system is compromised in case of a compromise; therefore, the system is weak to address the APTs.

Albanese et al. (2015) proposed an approach to defeat an attacker’s fingerprinting effort through deception. They manipulated outgoing traffic to resemble traffic generated by a host with a different operating system to defeat OS fingerprinting. Similarly, they modified the service banner by intercepting and manipulating certain packets before leaving the host or network to defeat service fingerprinting. Their approach showed that it could efficiently and effectively deceive an attacker.

Jajodia et al. (2019) described recent advancements in cybersecurity, including the use of Adaption Techniques (AT). Adaption Techniques includes Moving Target Defense (MTD) and Adaptive Cyber Defense (ACD) strategies. AT focuses on producing uncertainty within a network, endpoint, or application. The purpose of MTD is to dynamically change the cyber environment to increase the difficulty for the threat actor.

Ayoade et al. (2020) stated by 2022, defensive deception products are expected to be a $2 billion industry. One challenge the industries encounter is the difficulty in conducting human subject experiments. Unlike other cyber defensive tools such as anti-malware, defensive cyber deception technologies are designed to interact with humans due to the difficulty and time-intensive requirements to assess the effectiveness of cyber defensive deception technologies. Existing studies on the effectiveness of cyber defensive deception technologies are ambiguous (Ayoade et al., 2020).

In conclusion, simulation of user activities and the network behaviour in the lifelike decoys is a task that needs careful consideration by considering the skills owned by the attackers and the extent to which they can go into searching for the facts from decoys have.

2.10. Research Gaps

There is limited research that focuses on improving the cybersecurity tools and techniques used in the institutions of higher learning in Kenya (Serianu, 2020). It means the cyber attacks and threats, though ever-rising, is unnoticed. The institutions risk losing data and finances due to inappropriate measures to secure core systems and infrastructure (Chetalam, 2018).

Also, the decoy frameworks lack high deceptive power due to the limited generation of believable user activities and network traffic, which are vital in misinforming the attackers (Sun et al., 2019). It makes the decoys easily identifiable from the existing systems due to their behaviour and characteristics.

We are combining the efforts done by Sun et al. (2019) and Updyke et al. (2018). This research work will address the shortcomings of the previous works by incorporating the

General HOSTS (GHOSTS) framework (Updyke et al., 2018) in developing a scalable hybrid decoy framework with adequate user believable activities and network traffic generation, which lacks in other frameworks.

CHAPTER THREE

RESEARCH METHODOLOGY

3.1. Introduction

The proposed framework is a decoy system composed of intelligent decoys that can effectively misinform the attackers into believing they have accessed the natural system by using two layers of decoys and intercepting the commands from the attacker while directing them to a safer server decoy. The attacks are logged and analyzed using elastic search, which can inform cyber experts in real-time.

3.2. Cyber threats analysis and evasion techniques

3.2.1. Research Design

To effectively obtain desired output, this research work adopted the definition by Babbie et al. (2011), where they defined the design of studies as a study plan. The study design also includes a comprehensive plan to undertake research. For this study, a descriptive research design was adopted. Descriptive research was an appropriate choice because the research aim is to identify characteristics, frequencies, trends, and categories (Bradshaw et al., 2017). It was helpful in this study because there was scanty information regarding tools and techniques used to thwart cyberattacks in higher learning institutions, which builds the research's core purpose.

3.2.2. Area of the Study

This research study was conducted in Kenya's sampled universities and constituent colleges because they have elaborate systems and network infrastructure (Mwathi, 2018). According to Commission for University Education (CUE), there are 31 public and 18 private universities, 11 constituent colleges, and 13 with letters of interim authority forming 73 institutions. However, the researchers used a sample of universities (Mwathi, 2018). Sample derivation was done using Slovin's Formula (Tejada & Punzalan, 2012) was used to calculate the sample size (n) given the population size (N) and a margin of error (e). It was computed as $n = N / (1+Ne^2)$

Whereas n = no.of samples,N = total population, e = margin of error. Since N= 73 and e = 0.1 therefore; n = 73 / (1 + 73 * 0.1^2) = 73 / (1 + 0.73) = 73/1.73 = 42

3.2.3. Target Population

To derive the target respondents, purposively selected ICT director/Head ICT and system security analyst/Network administrator in each of the sampled institutions of higher learning in Kenya in the survey. The reason for using purposive sampling is due to the assumption that heads of ICT departments and cybersecurity professionals know the existing tools used in securing networks and systems in the institutions. The target respondents were derived from the sample size above by multiplying the sample size n by two since each sample university required at least two respondents. Therefore, the sample respondents were; n x 2, of which n=42 hence; 42x2 = 84 respondents.

3.2.4. Sampling Techniques and Procedures

Purposive sampling was used to pick samples from this research. The concept was to select the sample according to some criterion deemed necessary for the research question (Etikan et al., 2016). This study used purposive sampling to collect sample participants in each sampled university because cybersecurity is being practised by a few IT professionals who have adequate knowledge of the network and systems security mechanisms(Guarte & Barrios, 2006).

3.2.5. Data Collection Methods

Primary data were collected using online questionnaires. Questionnaires were designed based on the specific objective to assess cybersecurity threats analysis and evasion techniques used in higher learning Institutions in Kenya and sent online via emails in the form of google forms due to COVID 19 guidelines. Secondary data were collected using a literature review. The data collection period was done between the 15th October 2020 and 17th January 2021.

3.2.6. Data Analysis

Quantitative data analysis was used and analyzed using the Statistical Package for Social Sciences (SPSS). Tables and figures were used to present the significant findings of the study. Quantitative data were derived from the questionnaires administered in the subject matter.

3.2.7. Validity and Reliability

The validity of the data collection tools, this research focused on content validity. Content validity assesses whether a test represents all construct aspects (Beck & Gable, 2001). Five experts in the cybersecurity domain tested the tools to ensure they were relevant and covered all aspects of information security in the institutions to ensure accuracy. Reliability, on the other hand, refers to how consistent a method is in measuring something. This research was measured by checking the internal consistency Cronbach alpha coefficient to test the reliability of which a coefficient of 0.73 was obtained. It is within the acceptable level for a research study, according to (Taber, 2018).

3.3. Developing an adaptive hybrid deception decoy framework

3.3.1. Prototype design

Due to stability and security, the framework prototype was built on the Linux operating system, most institutions' widely used server operating system. The prototype encompassed two layers of the decoys, the lightweight front-end decoys installed in Linux LXCs and the backend systems installed on back-end LXCs. To interlink the two layers, this adopted research work used the Honssh tunnel. The LXCs were used due to their fewer costs on resources and the ability to scale the system up. The entire system was integrated on a single host machine running Proxmox VE 6.1. The host computer features were; Core i5 Intel processor and 16 Gb memory with 500 Gb storage. Three virtual machines were created for logging and analysis and another for gateway services over HonSSH; the third is an attacker VM installed with a Kali Linux operating system.

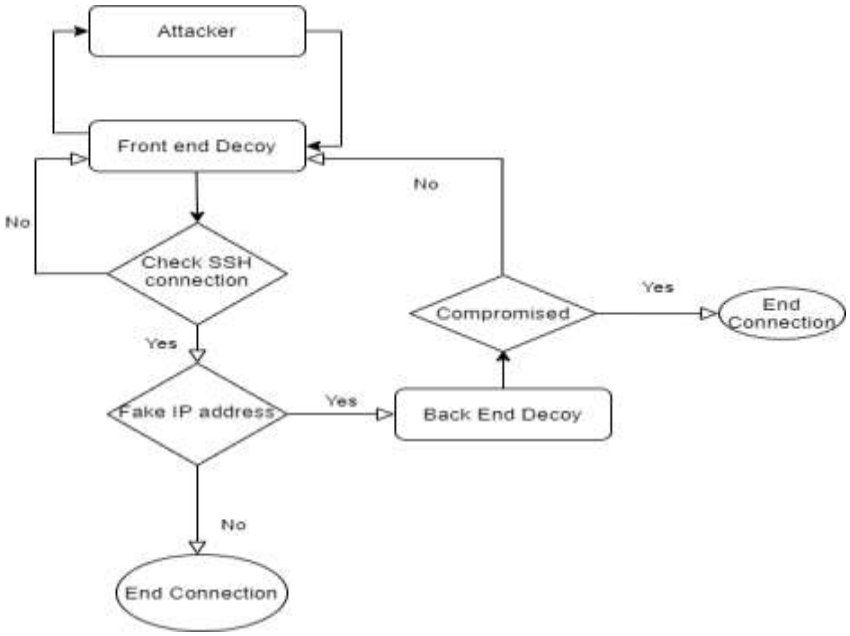


Figure 3.1: Decoy Framework Flowchart

3.3.2. Decoy system setup

The LXCs were set up in the Proxmox VE with two different setups, one for the front end and the other for the back end. A Linux VM was set up for Honssh that creates two separate channels between the two decoys. A separate Logstash server was installed to log and analyse the attacks with Elasticsearch and Grafana on Ubuntu server 18.0 LTS. Grafana was used to visualizing the logs. Both of these tools are based on Elasticsearch, which is used for storing logs. An attacker VM was installed with Kali Linux, where sample attacks were launched on the front-end decoys which are exposed. The logical design of the framework is outlined in Figure 3.2 below.

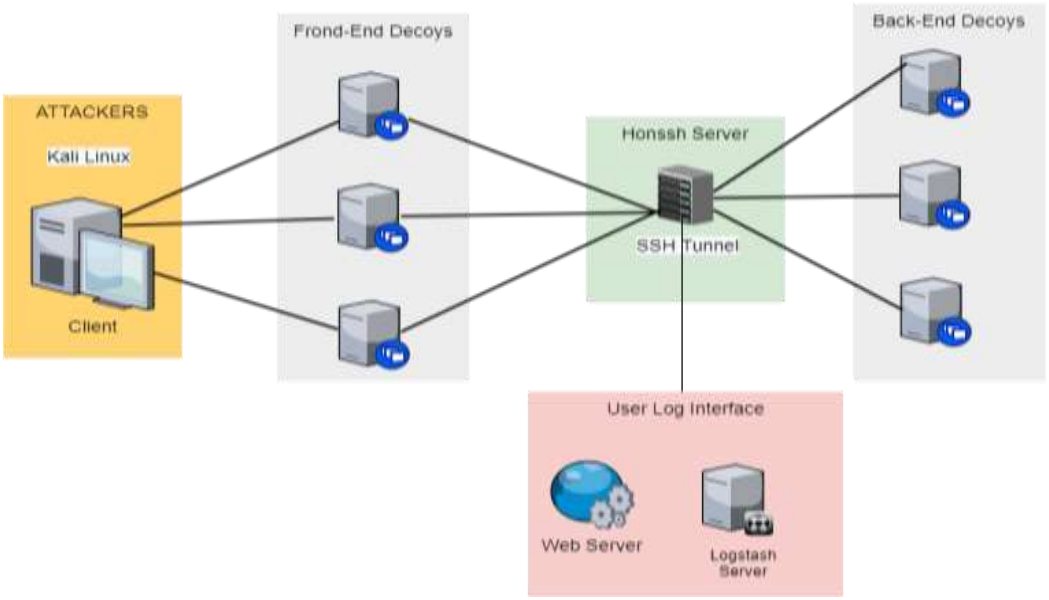


Figure 3.2:Decoy framework setup

3.4. Front-end server

The front-end server was installed on Ubuntu server 18.0 LTS over Proxmox LXC container. The SSH service was installed without any other service running on it. It is to make sure that the attacker's commands are captured and redirected to the backend decoy. To ensure the effectiveness of deception, the front-end IP address was bound to a fake virtual IP address in the HonSSH. It ensures the decoy server is not identifiable while allowing the command and execution in the back-end decoys. Figure 3.3 below illustrates the arrangement of the decoy servers and the HonSSH server.

Figure 3.3: Integration of the front-end and back-end decoys

3.5. Back-end servers

The back-end decoy server was installed on Proxmox LXC with similar attributes as the corresponding front-end decoy, Ubuntu server 18.0 LTS. Also, other services like SSH, Apache, MySQL, and SMTP were installed to mimic actual server services. The Honssh fake virtual IP was bound with the back-end decoy IP address, enhancing

transparent command interception and redirection. The fake IP addresses are generated randomly, effectively hiding the connection between the two decoy layers.

3.6. Network traffic and user activities simulation

This project used the General HOSTS framework to simulate the network traffic generation and user activities in the decoys, creating realistic network traffic in the form of context-driven user activity on a network. The backend decoy had GHOSTS installed but worked so that the attackers would not directly relate the activities to any service with the decoy's operating systems. The Ghost server was installed separately to monitor the performance of the agent in the decoy. The information is pushed to the Grafana front-end in the ghost server for easy retrieval and visualizing.

3.7. Data Collection

Data were collected using Logstash, an open-source data collection engine with real-time pipelining capabilities. Logstash can dynamically unify data from disparate sources and normalize the data into destinations. It cleanses and democratizes data for diverse advanced downstream analytics and visualization use. Figure 3.4 below shows the data collection process from the Honssh server to display in the portal.

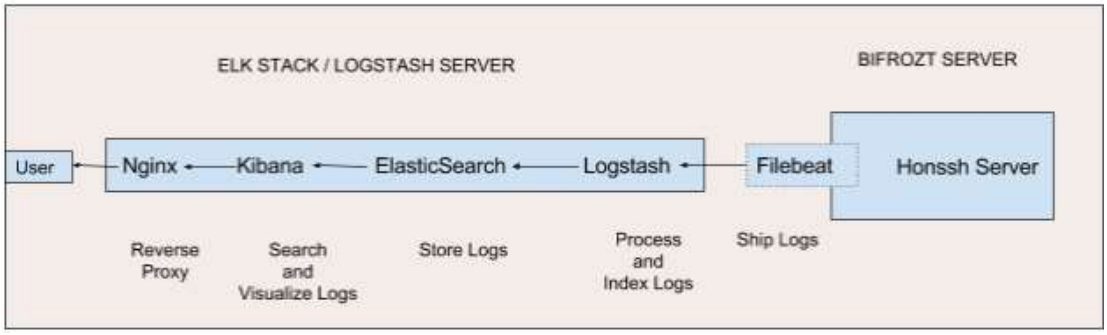


Figure 3.4: Data collection from the decoys

Filebeat was installed in the Honssh Server to ship logs from the server to Logstash Server. The server process and indexes the Logs which Elasticsearch then stores. Kibana is used as the interface for visualizing and searching for the stored logs. Users access the interface through Kibana, which is powered by a reverse proxy web server by Nginx. Other more filtered logs in the Honssh server are pushed to Maria DB SQL

server, which in turn a front-end system can be developed for a real-time view of the logs.

3.8. Effectiveness and efficiency of the adaptive hybrid decoy system

Three metrics were measured: latency, scalability, and throughput to test the decoy system's performance because the research was done by scholars like Sun et al. (2019) and Sun et al. (2017) focused on these metrics in measuring the effectiveness and efficiency.

The latency was measured by checking the times of running remotely issued attacker commands in the front-end decoy and the respective response by the back-end decoy.

The scalability was measured by increasing the number of LXC containers with installed decoys with respective containers. Simultaneous booting of the LXC based decoys and their time to boot and start respective services while checking the overall performance when the number of decoys increased. The effectiveness of setup was recommended for redundancy and alternative decoy connection when one of the decoy crushes or is adversely affected by the nature of the attacks.

The throughput of the decoys was measured by checking the rate at which data flows from the decoy server to the attacker machine. It informed how efficient the setup is in luring the attacker and obtaining necessary information from the server.

CHAPTER FOUR

RESULTS

4.1. Introduction

This chapter presents the findings obtained from the survey conducted on tools and techniques used by the institutions of higher learning in Kenya, the decoy framework development and testing of the efficiency and effectiveness of the framework, discussion, and the contribution of deception decoys in cybersecurity.

4.2. Cyber threats, tools and evasion techniques

The respondents of this survey were 84 ICT practitioners from chartered universities and university colleges in Kenya, out of which 67 responded to the questionnaires representing a 79.8% response rate. This response rate is sufficient for research, according to Mugenda and Mugenda (2003). Figure 4.1 below outlines the number of ICT personnel that work on cybersecurity in their universities.

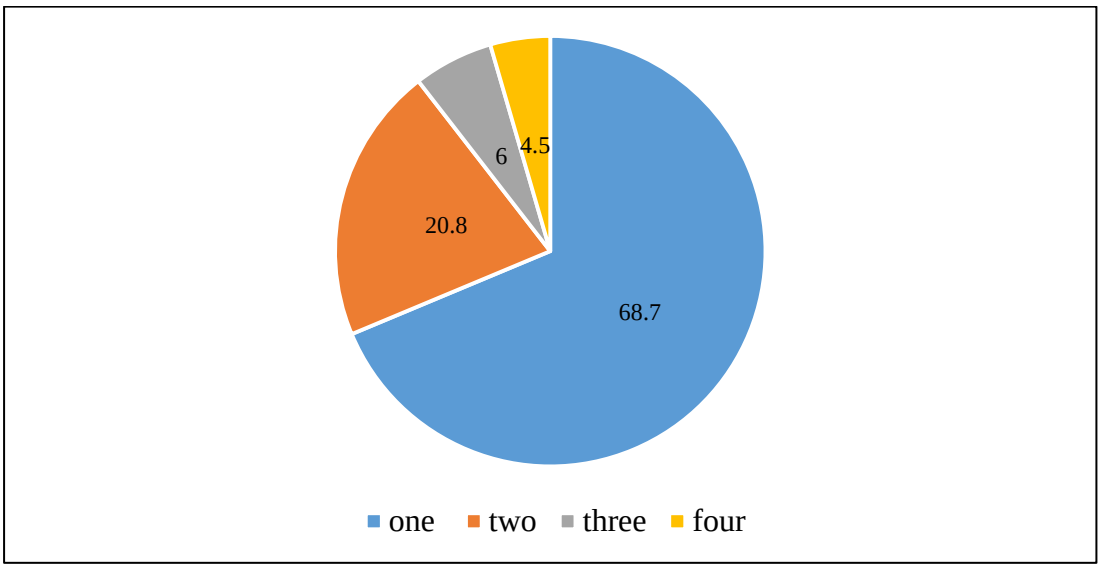


Figure 4.1:Number of personnel involved in cybersecurity

All the respondents stated that they have cybersecurity and IT infrastructure in their institutions. All of them are aware of the security features employed by the universities in securing the systems and networks. In addition, all the institutions in the study had

at least one IT professional dedicated to cybersecurity (68.7% had one, 20.8% had two, 6.0% had three, and 4.5% had four). It means that the institutions have put in place structures to mitigate cyber threats. However, human resource capacity is still lower compared to the nature of work the personnel need to undertake at any given time. Therefore, the accuracy of the cyber reports is inefficient in the institutions due to both capacity and infrastructure. Most universities and colleges have cybersecurity policies with 76.1% in place, while 23% lack policy establishment. However, the guidelines are limited to the scope of the current tools used in the institutions. Though the respondents stated that the policy was available, it was engulfed in the ICT policy, which is one article rather than a stand-alone policy.

Additionally, 79.1% have well outlined cyber programs as the rest, 20.9%, don't have any program in place, of which the 64.2% are being developed or configured by the in-house team. In comparison, 35.8% outsource the service to other experts to do it on their behalf. They're either trained on managing the systems or completely hiring service providers for installation and support.

The institutions outsourcing cybersecurity roles are a threat because the services depend on the service providers' integrity and goodwill. If the systems are installed internally, IT teams are reluctant to take up security matters more seriously due to confidence over-reliance on the skills and reports obtained from the service providers. It is expensive in the long run.

The number of systems and network infrastructure secured from cyber threats varies from one institution to another, with 59.7% with more than five systems, 32.8% with less than five systems under the infrastructure, and 7.5% with only five systems operating in the secured environment using the available tools. It means that a number of the systems are under an unsecured climate, and therefore, data collected from the cyber tools are purely inclined and biased on the system under their configured tools. It posed a significant challenge to the real cybersecurity position of the institutions.

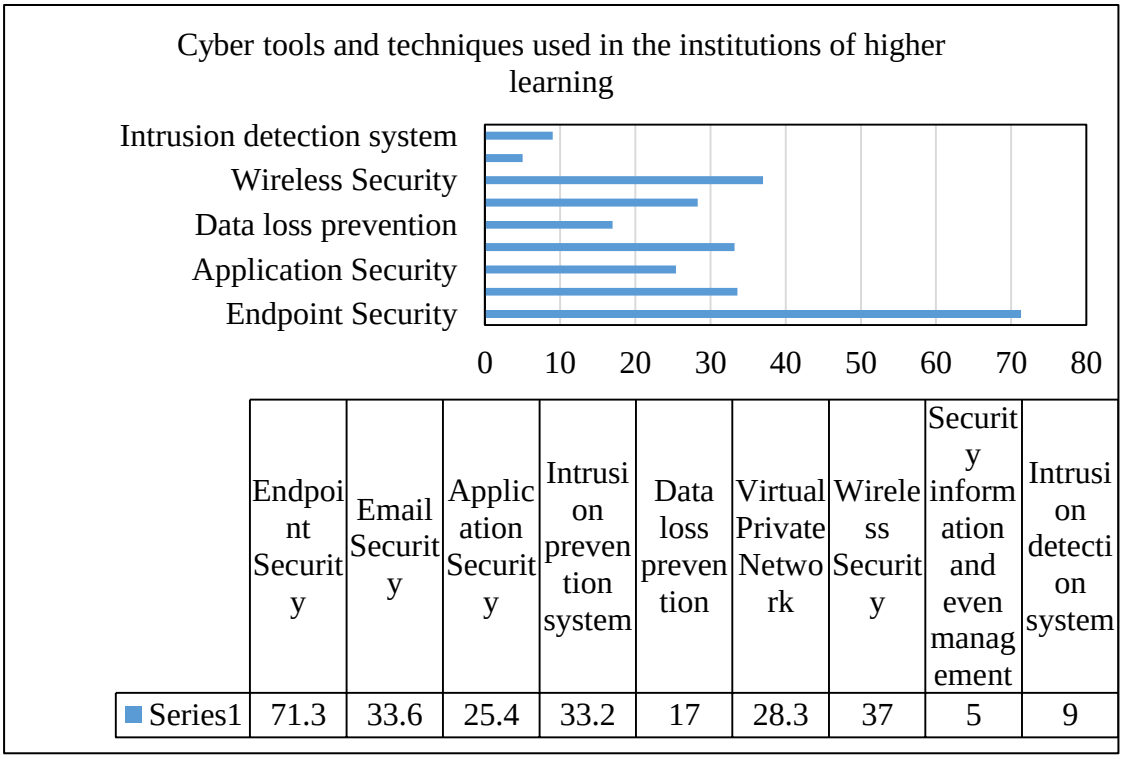
The most common systems in the institutions accessed via the web under the secured and unsecured environment are; Staff Portal, Student Portal, Koha, and E-learning

system. The other systems include; Help Desk, OPAC, Banking Integration APIs, Student Clearance, and Digital Repository.

Based on the responses from the practitioners concerning cybersecurity status in the institutions, the following are the tools used by the institutions; Firewalls, Access Control, Anti-Malware, Endpoint Security, Email Security, Application Security, Intrusion Prevention System, Intrusion Detection System, Wireless Security and Security information and event management system.

4.2.1. Cyber tools and techniques used in thwarting the threats

It was revealed that the institutions use various cyber tools in countering the identified cyber threats, of which 67.2% use more than eight multiple tools to secure their systems and networks while 32.8% use less than eight cyber tools in their infrastructure. All institutions use access control, firewalls, and anti-malware software in their devices within their network. 71.3% of the institutions use endpoint security, while 33.6% use email security, while 25.4% use application security in securing the applications in their network. Figure 4.2 has a detailed distribution of the tools used in the institutions.



	Endpoint Security	Email Security	Application Security	Intrusion prevention system	Data loss prevention	Virtual Private Network	Wireless Security	Security information and even management	Intrusion detection system
Series1	71.3	33.6	25.4	33.2	17	28.3	37	5	9

Figure 4.2:Cyber tools and techniques used in the institutions of higher learning

Intrusion prevention system is used by 33.2% of the institutions, while on the other hand, only 17% of them use the data loss prevention systems. A few institutions have adopted virtual private networks (28.3%) to secure end-to-end communication between internal systems and remote access to services via secure tunnels. Wireless security seems poorly done though most institutions have installed wireless local area networks, with 37% of them having implemented necessary security measures. Security information and event management (SIEM) and intrusion detection system have least been implemented with 5% and 9%, respectively.

The tools used by the institutions in cybersecurity have 85.1% term them as applicable while 14.9% feel that the systems are not effective enough to thwart the cyber attacks. 65.7% of the practitioners understand how the systems work, while 34.3% lack technical skills to manage or operate the existing systems. Most of the cyber tools used by the practitioners in their institution are open source (76.1%), while the rest are either outsourced or subscribed. However, there is are a general feeling that the systems need improvement because the team lack a proper way of tracking attacks. Especially the stealthy ones with 88.1% of them suggest that the tools need improvement to be more effective and efficient in securing the systems and the network they manage.

The fact that the tools used to give the ICT personnel confidence of service is not a true reflection of the quality of the cybersecurity in place because of the types of tools used. The basic knowledge is that no significant attack was recorded if anyone was interested in harming their systems. Another contributing factor of the biased response is the level of skills the ICT personnel have concerning cybersecurity. They assume that the hackers target the banks and financial institutions for financial gain and educational institutions are less of their interest. In a real sense, they have much information in academia and research and are actively developing the systems without security in mind.

4.2.2. Common cyber threats in the institutions

Malware and phishing are common cyber threats across universities and colleges, while 84.1% of them face denial of service attacks, as 45.3% suffer from the man-in-the-middle attack (Figure 4.3). The attacks were launched on the systems within and outside the institutions. The attacks noticed were captured mainly by the firewalls and IDS or

IPS in the institutions. Other tools that the institutions used to detect the attacks were; SIEMS (18%), Honeypots (12%), and decoys (5%).

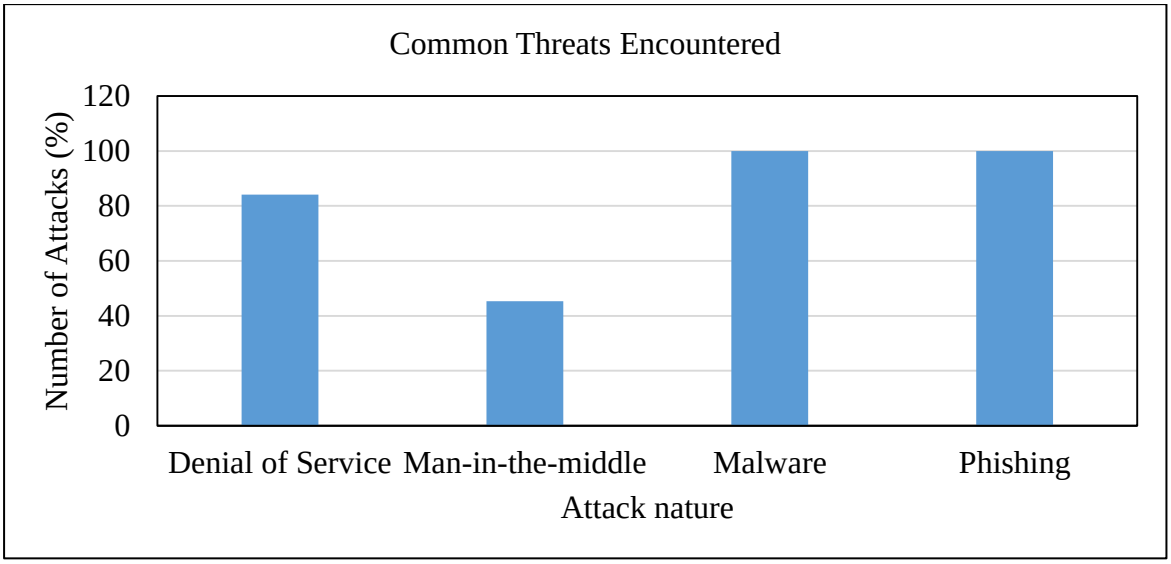


Figure 4.3:Common Threats in the institutions

Computer malware remains one of the threats faced by users in most organisations basically due to lack of antivirus or ignoring the messages the anti-virus prompts them to scan or prevent malicious software from running. In most cases, cracked software is used every day by many users and even IT professionals. It makes the anti-virus irrelevant hence making the machines prone to malware attacks. Other sources of malware include torrents files, fake emails, and corrupted flash disks. According to Serianu limited (Serianu, 2020), a pan-African cybersecurity firm based in Nairobi, Kenya, in the year 2020, most organisations recorded the highest cyberattacks ever (Figure 4.4). It is attributed to COVID 19 protocol especially working from home, which found the IT teams unprepared to deal with attacks based on the new model of a user operation from home.

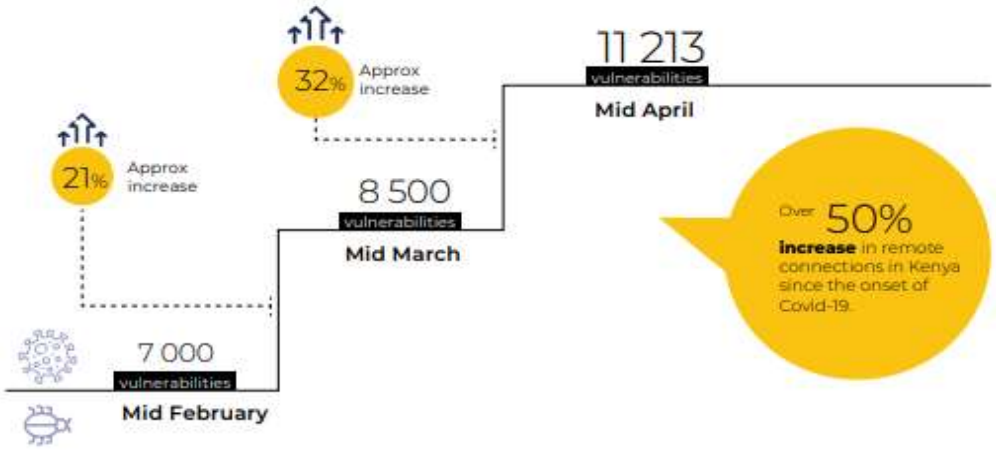


Figure 4.4:Remote connection Vulnerabilities in Kenya (Serianu,2020)

The report by Serianu shows a similar trend in the learning institutions in Kenya. Further in their report, phishing, malware, and exploitation of the telework infrastructure attack were witnessed more in 2020. From our studies, malware distribution and phishing were reported as the common cyber attacks in the institutions. It shows that a lot of work still needs to be done to ensure advanced security in the institutions of higher learning in Kenya to reduce the attacks as much as possible.

In the past 12 months, the institutions recorded various cyber cases depending on the detection. Most institutions recorded below 100 incidents (57%), while others recorded over 500 incidents (12%), 3% recorded between 100 and 500 incidents. In contrast, 18% did not record any cyber-related cases either by tools not being efficient or when no one monitoring them (figure 4.5). All the cybercriminals targeted the Enterprise resource planning systems with associated service portals. Other systems are; payroll management systems (60%), websites (22%), and network or network devices (18%). The system security managers record the cyber-related cases (60%), and some (40%) lack proper database management on cyber issues for study and scrutiny.

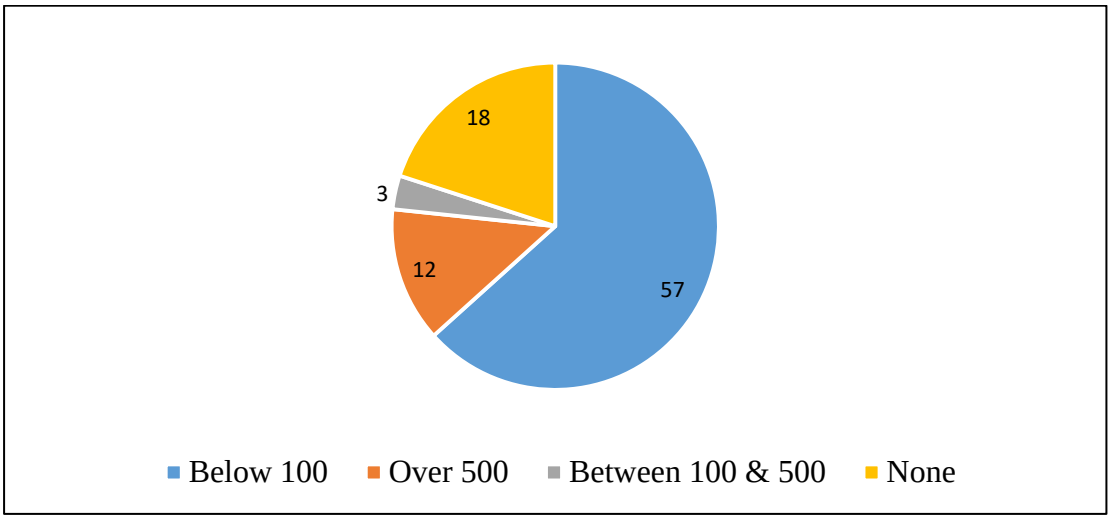


Figure 4.5:Annual record of cyber attacks

Most practitioners lack proper knowledge on decoys and deceptive technology though 51.9% can define the meaning very well while the rest (47.1%) lack knowledge and definition.

The institutions' annual records on cybersecurity-related incidents may not be accurate due to limitations outlined above, including the use of outdated tools, lack of real-time monitoring of the security tools, and lack of a skilled workforce to conduct proper cybersecurity analysis. The statistics show that attacks occur in the institutions and can even be more if well monitored. The attacks cannot be ignored even if it is because the growing interest in hackers deploying distributed denial of service, ransomware, and botnets is an alarming trend. They are considering the level of intellectual property in the institutions of higher learning, especially research findings and copyright details.

According to Serianu (Serianu, 2020), academia and research are the second (16%) after the manufacturing industry (37%) in the number of publicly accessible ports in the network or servers. It makes them prone to attacks by hackers (figure 4.6). The hackers conduct surveillance of the networks to determine the weak points then launch a heavy attack on the vulnerable network and services.

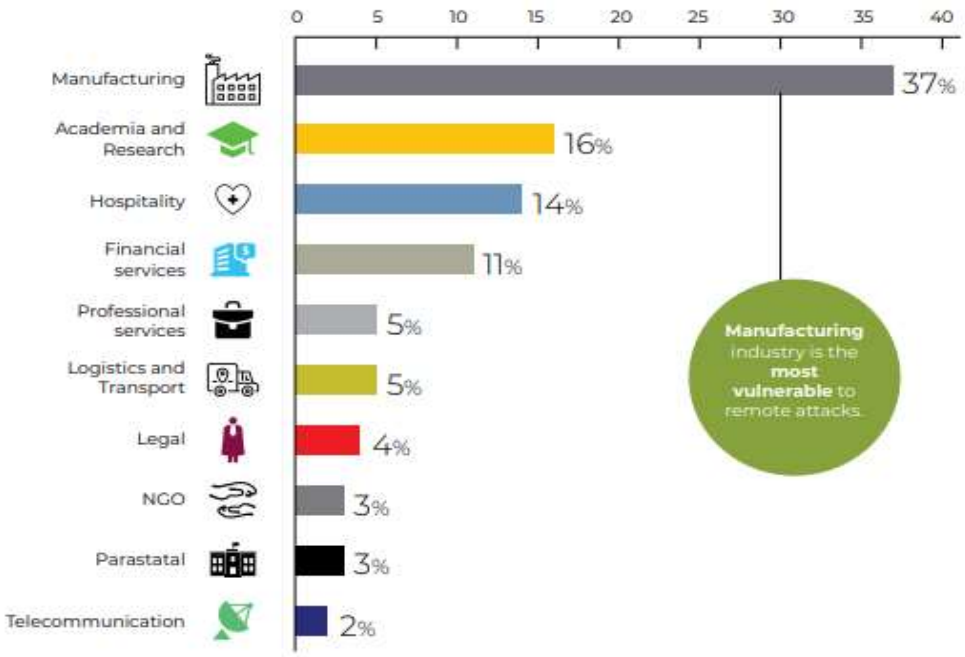


Figure 4.6:Publicly accessible ports in Kenya (Serianu, 2020)

The level of attacks witnessed shows that the institutions need to take cybersecurity protocols and measures to prevent fraud, intellectual property loss, financial property, and possible damage to infrastructure or delay of services.

The public accessible ports vary in the level of exploitation. For example, telnet protocol is the most vulnerable because it sends data in plain text format(Stahnke, 2006). The accessible ports in Kenya are outlined below in Figure 4.7. The meaning of this is that the level of cybersecurity in the learning institutions requires more elaborate mechanisms to deal with the vulnerabilities associated with the ports. It includes empowering the ICT teams, investing in more robust and up-to-date systems, conducting regular and sufficient cyber penetration tests, and patching the vulnerabilities as soon as they are discovered. The fact that telnet ports are very active shows that the administrators of the networks and the systems understand more diminutive of the secure channels in accessing the servers and networks.

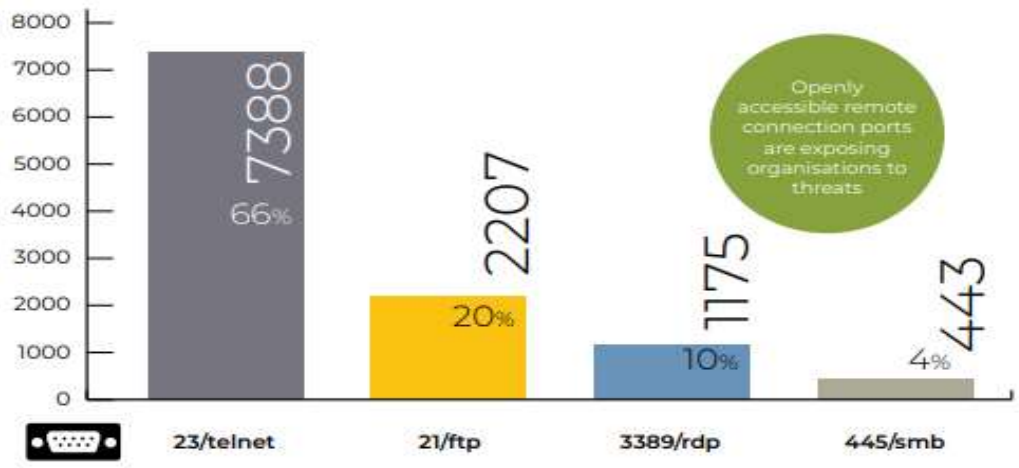


Figure 4.7:publicly Accessible ports in Kenya (Serianu, 2020)

The FTP port is used for file transfer; these findings show files and folders are shared over public infrastructure. There are other safe and secure alternatives to use and administer, like document management systems or cloud platforms in file sharing. The remote desktop access protocols should be minimized. The users are encouraged to use more secure alternatives like the Anydesk or Teamviewer software, which are relatively easier to use and convenient for any user to operate.

4.3. Deceptive decoy framework

4.3.1. Performance Evaluation

The decoy framework succeeded in an interception and redirecting all the commands entered from the attacker VM and responding to the respective queries as if it were a real server. The Honssh server generated random fake IPs to shield the decoys server while giving adequate responses as if it were a real server (Table 4.1).

Table 4.1:Sample commands entered in the decoys

	Command	Description	Attacker	Back End decoy
1	who	Logged in users at the time of the tests	edwin	edwin
2	uname -o	Check the running operating system	GNU/Linux	GNU/Linux
3	pwd	Print the working directory	/home/edwin	/home/edwin
4	ifconfig	Display the IP address	192.168.100.104	192.168.100.110
5	df –h /dev/sda1	Checking the disk space usage in partition dev/sda1	dev/sda1 236M 80M 144M 36% /boot	dev/sda1 236M 80M 144M 36% /boot
6	hostname	The name of the host/network	megasoft	megasoft

The table above shows that the decoy system managed to hide the actual IP address (**192.168.100.110**) of the back-end decoy, and this is because the Honssh fake IP address managed to prevent the real identity of both IP tunnels. Apart from that, all the other information queried by the attacker matches the exact information from the real back-end decoy server. The IP address that appears on the attacker shell shows the front-end decoy IP address (**192.168.100.104**). It means the attacker will assume that the front-end decoy has been successfully attacked, but what the attackers do is being monitored absolutely. The screenshot below (figure 4.8) shows the IP address in the decoy.

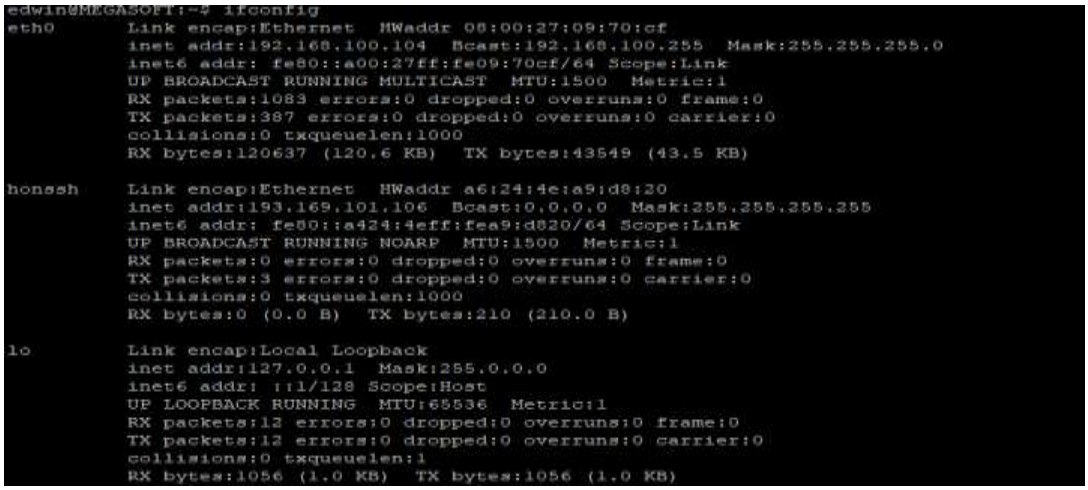


Figure 4.8:Screenshot of the front-end decoy network address

The screenshot above shows the actual IP address of the front-end decoy (**192.168.100.104**) and the Honssh fake IP address (**193.169.101.106**) that bound two separate SSH tunnels. The Attacker will not notice the fake IP address.

To protect access to the Honssh server via a password-based attack, we introduced key-based authentication for security and moved the SSH server to a different port. This ensured port 22 is redirected to port 2220, which is the listening port in the Honssh, therefore, magnifying the aspect of command transparency and redirection.

4.3.2. Logs and Monitoring

To monitor the attacker's behaviour, we adopted real-time logs capture to a MySQL server. Also, we installed Filebeat, which pushed the results to the Logstash server for processing and visualization. The raw tables are shown below. The sampled attacks were performed between date 22nd May to 28th May 2021 (figure 4.9).

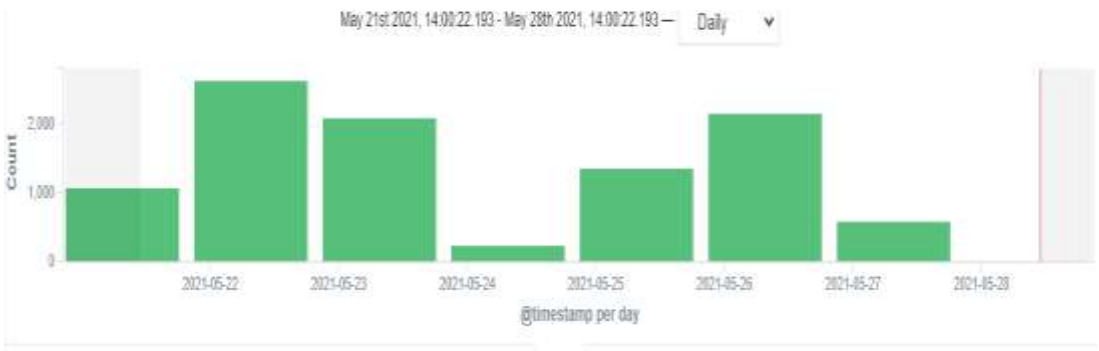


Figure 4.9: Logs from the back-end decoy

1	edwin	@sk.compiso	2021-05-22 21:19:31
1	edwin	@sk.compiso	2021-05-22 21:24:07
1	edwin	123456	2021-05-24 10:16:03
0	megasoft	123456	2021-05-24 14:30:12
0	megasoft	123456	2021-05-24 14:31:53

Figure 4.10: Database of the user login in the decoy framework

The decoy allowed user login when the real password was spoofed or with a near-real password percentage of 25%. The system also could detect a user login and capture the password used in plain text and the timestamp of the tried login.

2021-05-26 06:03:24	e65b996143e34621955c62d73fed3f2b	□hostname
2021-05-26 06:03:32	e65b996143e34621955c62d73fed3f2b	^C
2021-05-27 19:03:34	2b7312d95b844cc6a5c646445cd0f3b3	ifconfig
2021-05-27 19:03:42	2b7312d95b844cc6a5c646445cd0f3b3	who
2021-05-27 19:06:15	2b7312d95b844cc6a5c646445cd0f3b3	ip addr show eth0

Figure 4.11:Commands captured from attackers shell

The decoy framework can capture accurately the commands entered and the attacker's IP and the SSH client used during the attack. They managed to send emails on the connection and disconnection of the decoy server. The system generated the list of attacker IPs, the time they started to attack, and when they left the decoys. These details are vital in facilitating cybersecurity specialists to secure their systems by learning from the attacker.

4.3.3. Network traffic and user activities simulation

The clients used in the back-end decoy generate believable network traffic using the Ghosts framework. The way the network is busy and sometimes the client machines generate adequate user-like activities in the decoy give proof that the decoy framework can lure the attacker into the decoys believing they are access real system.

Ghost Framework uses realistic non-player character (NPC) orchestration, which means that it mimics real system users in creating documents, issuing commands, sending emails, and other functions. In our approach, the ghost client agent was installed in the back-end decoy to mimic the user **edwin**. A ghost server was established

to monitor the agents, performance and determine if realistic services were created To monitor the functions.

Figure 4.12:Decoy non-player character (NPC) activities

The screenshot above shows the activities in the back-end decoy over 28 days. The green part covers the timeline at a maximum of 8 on 13th May 2021. Minimum of 0 for the days the decoy was offline on 6th May 2021. The decoy health was ok apart from one day when it detected malicious activity. The polls were at the maximum of 36 and 0 on the day the ghost agent was offline. The total activities recorded for 28 days were 240,612, which are substantial enough to convince an attacker of user activity in the decoy. The day the ghost agent was offline shows that the decoy activities are influenced by the decoy and not the decoy itself.

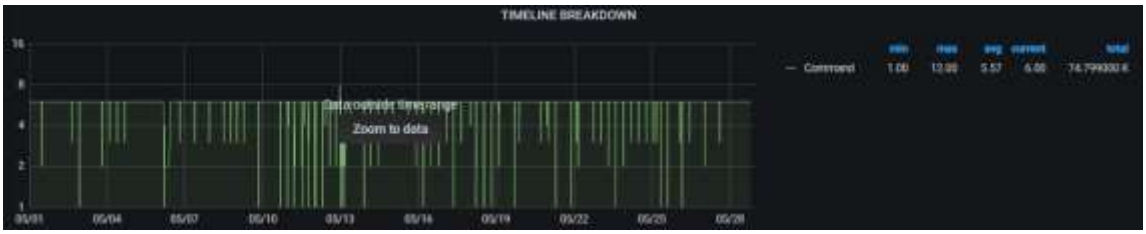


Figure 4.13:Commands tracking from the decoy framework

The total commands issued were 74,799, with an average of 5.57commands per day and a maximum of 12 commands issued on 13th May 2021. A minimum of one command was recorded on 5th May 2021.

4.4. Effectiveness and efficiency of the adaptive hybrid decoy system

4.4.1. Latency

In this study, commands latency was tested through execution in the real server and remote decoy over the intercepted shell, and the results are as follows (Table 4.2);

Table 4.2:Latency on the execution of remote commands in the decoy

Commands	Redirected	Real server
Uname	real 0m0.002s	real 0m0.004s
Hostname	real 0m0.003s	real 0m0.004s
Ls	real 0m0.002s	real 0m0.004s
Ifconfig	real 0m0.008s	real 0m0.013s
Who	real 0m0.004s	real 0m0.011s
Pwd	real 0m0.000s	real 0m0.003s

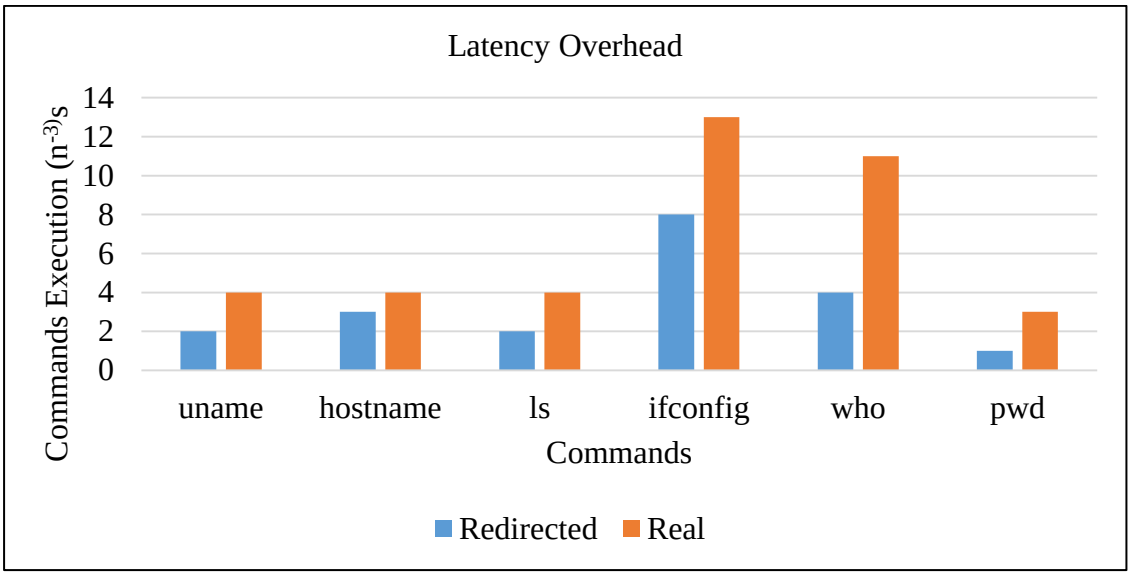


Figure 4.14:Graphical representation of the execution of the command

The above commands established that the decoy framework could intercept and redirect the commands with less latency overhead. The intercepted commands are executed typically just like the real server interaction with a slightly lower execution time. The explains why the use of layered decoys is a good approach. It is not easy for attackers to notice the difference between the actual servers and the decoys, therefore, boosting the aspect of deception through similarity of services.

(Vrable et al., 2005) developed a similar decoy framework and the execution of commands took a relatively more extended amount of time than the framework created in this study. The figure below shows their output. From the studies done, the framework is more effective than Vrable developed due to double deception enhanced by the fake IP address that binds the decoys into one machine.

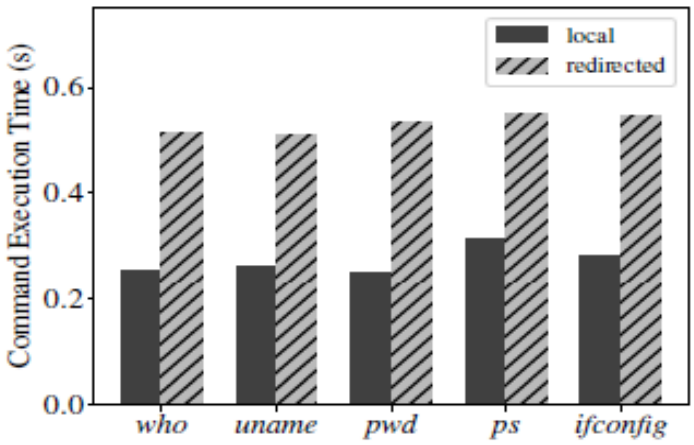


Figure 4.15: Latency Overhead of Command Redirection (Vrable et al. 2019)

4.4.2. Throughput

Throughput was measured between the backend decoy and the attacker to establish the rate at which data are relayed between the two VMs. The test was done using Jperf (figure 4.16), a graphical tool that simulates what Iperf does in real-time between two servers.

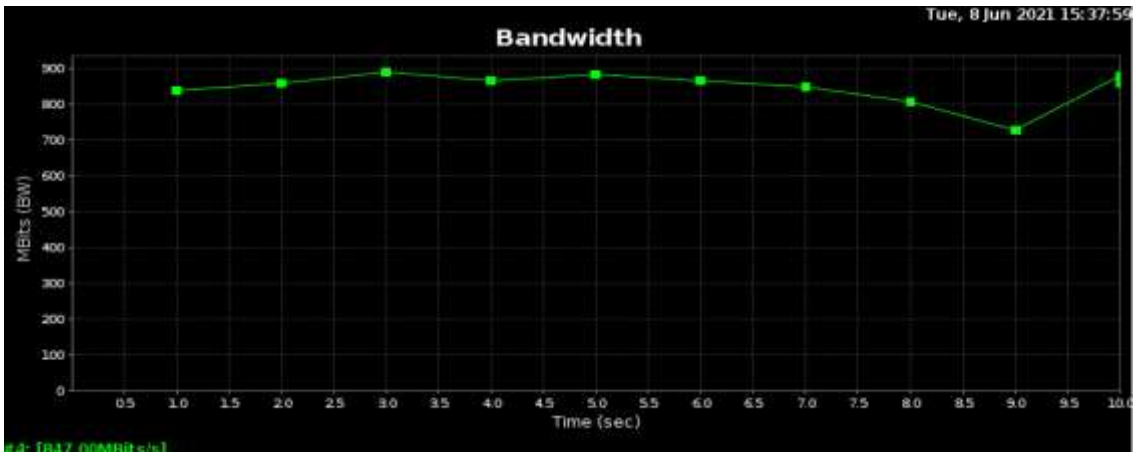


Figure 4.16:Throughput Measurement between the decoy and attacker

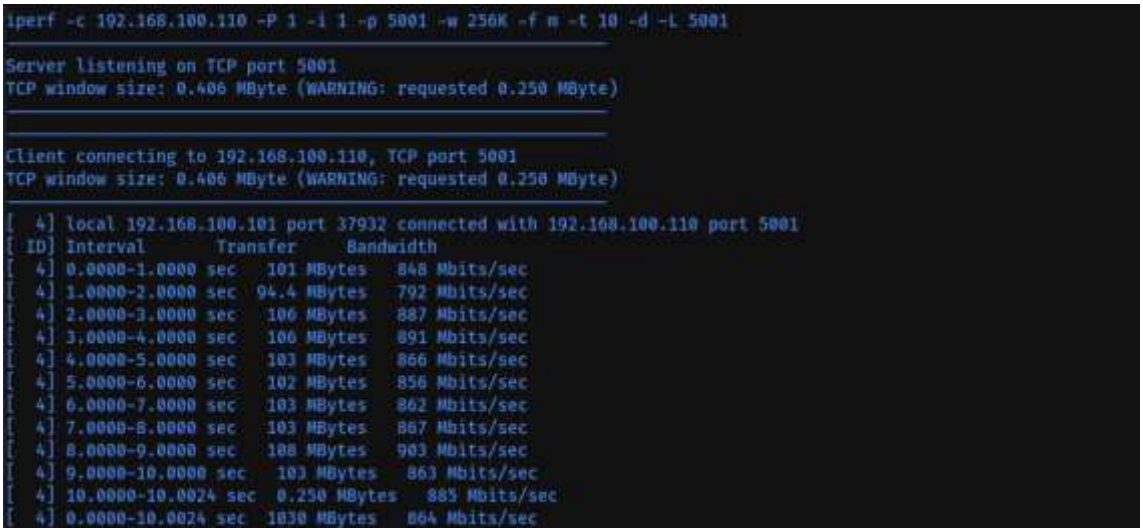


Figure 4.17:Iperf results in the decoy framework

From the results that were carried at an interval of 10 secs, a total of 1030 Mbs were transferred with a speed of 864Mbits/s. It is relatively a reasonable speed at which the system operates under normal operations. Therefore relaying the data between the decoy framework and the attacker is within acceptable speed limits to convince the attacker.

4.4.3. Scalability

We tested the Linux containers' speed to boot and load the necessary system services effectively to try how scalable the decoy framework can be. The table below (Table 4.3) shows the experiment done on six LXCs and six VMs.

Table 4.3:Boot speeds of the LXCs and VMs

	LXCs	**VMs**
1	12.108s	18.276s
2	7.909s	39.361s
3	5.681s	33.596s
4	6.543s	28.657s
5	7.485s	42.241s
6	8.328s	22.107s

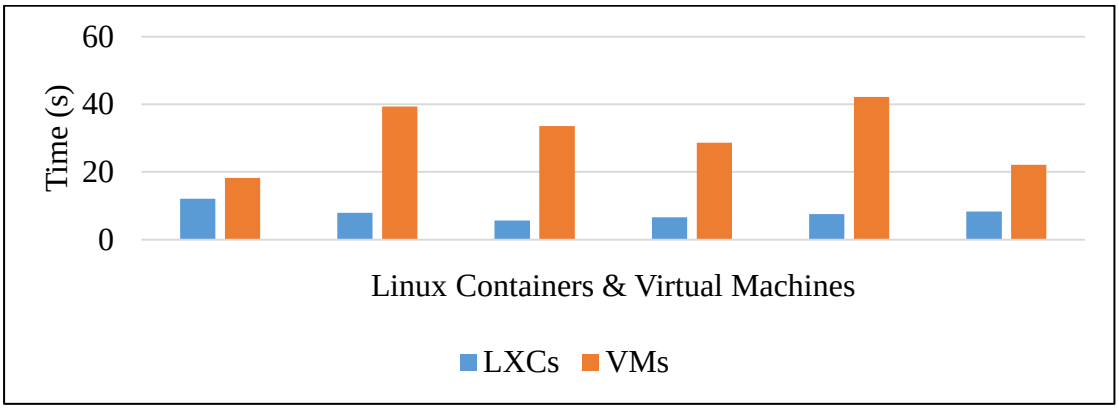


Figure 4.18:Graph showing the LXCs and VMs booting speed

The experiment showed (figure 4.18) that it is relatively faster to boot LXCs than respective virtual machines; therefore, it is easy to deploy and configure the LXCs in the decoy framework. It was also noted that the speed at which the commands respond depends on the command executed.

The other studies, such as the experiment done by Vrable et al. figure 4.19, show that the trend of booting time is similar. Though the speed is affected mainly by the type of environment deployed, the results from the studies show the LXCs have a relatively lower time to boot, which enhances scalability and efficiency.

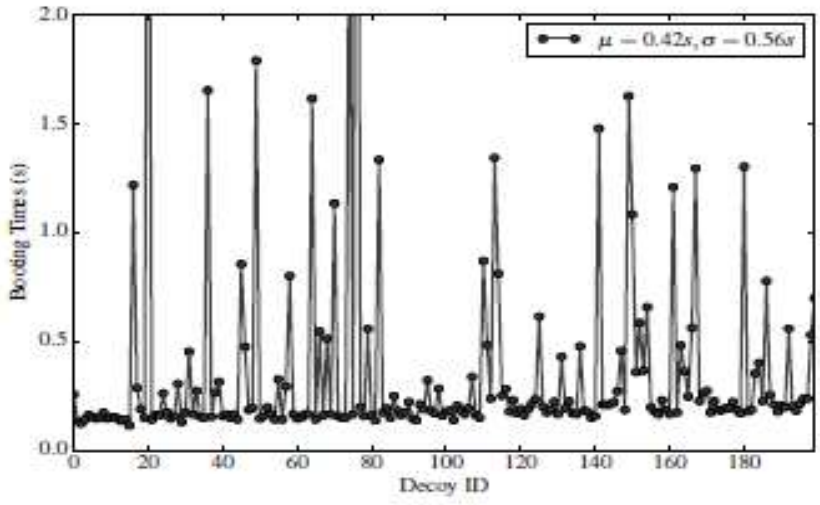


Figure 4.19: Scalability of the decoys (Vrable et al.)

Though the above figure 4.19 shows the trend in about 200 LXCs, the trend shows that they take less time to boot compared to virtual machines and even the host machines. The attributes provide flexibility of deployment and easy repayment if, by chance, one container is destroyed or compromised in the cyberattack session.

CHAPTER FIVE

DISCUSSION, CONCLUSION, AND RECOMMENDATIONS

5.1. Introduction

The findings outlined in this chapter are in connection with the formulated research questions. The chapter is divided into four segments: Introduction, discussion, conclusion, and recommendations.

5.2. Discussion

5.2.1. Cyber security tools and techniques used in the institutions

A firewall is the most used tool across institutions. And it works by packet filtering, allowing and denying traffic and packets from different networks (Hamed et al., 2006). Other firewalls use proxy services; essentially, creating a mirror of the computer behind the firewall prevents direct connections between the customer device and the incoming packets, protecting the network location from potential bad actors. The institutions use different types of firewalls, but most use PfSense, Squid, and Cisco ASA. A few of them have Sophos and Huawei firewalls. Though efficient in handling packets, the firewalls are stateless, meaning they don't inspect other elements apart from the packet headers therefore ineffective in preventing the attacks successfully (Nwanze & Summerville, 2008). A stateful firewall is a new type of firewall created to attend to the challenges encountered by stateless firewalls, though the ICT practitioners are yet to embrace them.

Wireless local area networks (WLAN) serve most of the users due to availability and reliability. However, though institutions have wireless networks in place (Mwathi, 2018), adequate security is still an issue, with 37% only managing to secure their networks. Passwords are easy to discover or crack. The connections are left unencrypted either due to unawareness on the side of the ICT administrators or the types of devices used to lack the security features (Masai & Wanja, 2016). The worst scenario is when some institutions have not implemented network segregation, rendering both office devices and other users like students to access the same services. Some of the

information is classified, hence bringing about an information security risk to the institutions.

Anti-malware software such as anti-virus and endpoint solutions serve as the primary protection of the client machines. These studies noted that up to 71.3% of the institutions had endpoint security in place. In most cases, installing the antivirus is the most preferred method of protecting users from unnecessary attacks. But the process is not 100% done, making the machines vulnerable and prone to malware, especially those the users download with emails or cracked software. Another problem is that the users hardly update their anti-virus software, making them unable to combat new and advanced attacks due to obsolete databases (Min & Varadharajan, 2016). Endpoint solution is a relatively better approach, but most institutions have not implemented the necessary infrastructure, especially the active directory. It makes it hard to administer updates on time and all the computers actively. In addition, the solutions are expensive to maintain because the institutions use corporate licenses to activate the software, especially Kaspersky and Sophos, which need annual subscriptions.

Security information and event management (SIEM) and intrusion detection systems (IDS) are primarily implemented together for better analysis of cyberattacks (Zeinali, 2016). SIEM conducts real-time system monitoring, notifies network admins of essential issues, and establishes event correlations. On the other hand, IDS is a device or software application that monitors a network or system for malicious activity and policy violations. Though necessary, less than 10% of the institutions have implemented these tools, thus a cyber risk. The system and network may be compromised, and no one monitors or knows what is going on in real-time. It is an excellent practice to implement such a system for cyber cases review and study the pattern of the attacks in any functional institution.

Implementation of virtual private network and application security in the institution recorded a lower percentage due to applicability. The users actively work on related financial applications like bank integration APIs and the institutions' ERP systems for faster transactions; therefore, they need to ensure a secure channel. The recent trend of remote access to services occasioned by working from home has encouraged VPN use.

However, about 28% of institutions have implemented VPN. The institutions rely mainly on Kenya Education Network Trust (KENET), their internet service provider, and manage their routers and firewalls at their level. In addition, they give the status of the services from time to time or need to use basis (Kashorda & Waema, 2014). Though effective, the level of security lies entirely on service providers, contrary to the organization's expectation of managing their infrastructure based on their security details and policies in place. It means that very few system security managers understand how to configure or manage the VPNs or application security though, in essence, they exist and work in their environment, and monitoring is in the form of periodic reports sent by the service providers, which are less effective

Institutions of higher learning in Kenya have less effective tools to thwart the rapidly building advance persistent threats. It comes when most universities are actively automating their services to suit the ever-rising information technology needs. The techniques in place lack agility, scalability, and adaptability of the new methods the cyber attackers adopt to distract, disrupt or encourage the service providers for financial gains. The use of firewalls, both stateful and stateless, is the most used approach to protect systems and networks in the universities and colleges, and other servers operating in the untrusted zones are entirely vulnerable due to lack of proper hardening protecting from attacks.

5.2.2. Adaptive cyber decoy

The cyber decoy implemented successfully relayed transparent commands from the attacker to the backend decoy through the HonSSH tunnel. Studies discovered that the logs could be monitored in real-time from the decoy to show the behaviour of the attacker is an active cyber attack. In addition, introducing a script to issue a fake IP address to the HonSSH server made the deception technique more robust and outlined the decoy system.

This research study outlines a method that is effective and less costly for the institutions. The proposed decoy framework uses the opensource tools which are readily available and customisable. Decoys are used to study and learn from the attackers and hence meant to inform and prevent further cyber exploitation and distractions. Effectiveness,

deception uses a combination of available security paradigms to securely divert or misinform the attackers and learn from them through logging their interactions with the decoys.

It was discovered that using deception is effective, especially by using transparent command redirection. The use of highly interactive honeypots can be used to develop hybrid decoys by combining the automating traffic generation and use of non-player character systems to influence the behaviour of the cyber decoys. Internet Address (IP) configurations in the decoys are essential; for example, in this research fake IP address was used to make the IP address similar both in the frontend and backend decoys which makes it complex to separate the layers of the decoys.

5.3. Conclusion

In conclusion, the systems and networks in the institutions of higher learning in Kenya are less effective in overcoming the Advanced Persistent Threats. There is scanty information about cyberattacks though they occur every second, and the level of awareness is a bit lower than expected of any academic institution. In this regard, this research proposed a cheaper and easier way to monitor, deceive, and deflect attackers from the actual system by combining deceptive decoys and complex networking techniques. The solution is meant to use readily available open-source tools to develop an adaptive, scalable decoy with a hybrid architecture of two layers of decoys, front end and back end decoys. Introducing these tools in the network will improve the existing tools in the institutions and help secure the systems. The use of information technology is gaining momentum in the universities and colleges in Kenya. Deceptive decoys use commonly known techniques to deceive the attackers using the systems that resemble the natural system by simulating everything and logging the interactions between the decoys and the attackers, preventing further escalation through hardening or tar pitting the systems.

Cyber deception is effective, especially when the decoys combine the faking of the user activities in the systems and advanced networking in the decoys. The extraction of the attacker's actions is vital because the attackers invent new ways to defeat already secured systems.

The experiments show that it is possible to develop effective decoy systems using open source resources by combining the already set system into a complex and effective cybersecurity tool. It was also discovered that it is possible to misdirect and misinform attackers into believing they have accessed an existing system while learning from them. The double logging of the events in the decoys with less overhead latency (average of 0.004s) commands execution time, which translates to efficient resource utilization and operation of the framework. In addition, it was noted that increasing user activities in the decoy system makes the decoy servers more realistic and convincing to the attacker. Hence, they can learn more due to confidence built by the system, with an average of 6.60 commands per day. The decoy framework exhibit scalability due to minimum resource utilization, giving an average of 8.008s boot time compared to virtual machines boot time average of 30.71 seconds.

5.4. Recommendations

This study recommends ICT practitioners use open-source decoys systems to protect the networks and systems. The decoys are hardened to prevent accidental discovery by the attackers in the active exploitation. Institutions of higher learning should beef up the network and security infrastructure to ensure protection against Advanced Persistent Threats (APTs) which are ever rising in the current information age. In addition, they should embrace decoys using readily available tools that can be customized and enhanced to come with practical tools both in cost and resource consumption.

Tar pitting is another way to ensure that the existing systems are further protected from the attacker by using delay and protect tactics. The defenders of the security systems should have the attacker's mind as they develop the adaptive decoys. It ensures that the systems designed are practical and efficient to run in the organisation. The use of LXCs in the decoys proved more effective due to their resource consumption and operation than using the VMs. It enables both scalability and agility and system operation in active cyber attack phenomena.

Based on the survey findings, it is recommended that the policymakers and stakeholders implement the cyber policies and ensure proper tools and techniques are used to protect the data and systems. Another area of research is integrating the decoys in the existing

cyber tools such as the firewalls, then check on their relative efficiency in combating the cyber attacks.

REFERENCES

Aggarwal, P., Gonzalez, C., & Dutt, V. (2016). Cyber-security: Role of deception in cyber-attack detection. In *Advances in human factors in cybersecurity* (pp. 85–96). Springer.

Alazab, A., Hobbs, M., Abawajy, J., & Alazab, M. (2012). Using feature selection for intrusion detection system. *2012 International Symposium on Communications and Information Technologies (ISCIT)*, 296–301.

Albanese, M., Battista, E., & Jajodia, S. (2015). A deception based approach for defeating OS and service fingerprinting. *2015 IEEE Conference on Communications and Network Security (CNS)*, 317–325.

Almeshekah, M. H., & Spafford, E. H. (2016). Cyber security deception. In *Cyber deception* (pp. 23–50). Springer.

Ayoade, G., Araujo, F., Al-Naami, K., Mustafa, A. M., Gao, Y., Hamlen, K. W., & Khan, L. (2020). Automating Cyberdeception Evaluation with Deep Learning. *HICSS*, 1–10.

Babbie, E., Mouton, J., & Strydom, H. (2011). The research process with reference to the research method section Social work theories and methodologies: Dubrovnik. *Croatia: North West University, Potchefstroom, South Africa Herman*.

Beck, C. T., & Gable, R. K. (2001). Ensuring content validity: An illustration of the process. *Journal of Nursing Measurement*, *9*(2), 201–215.

Beham, M., Vlad, M., & Reiser, H. P. (2013). Intrusion detection and honeypots in nested virtualization environments. *2013 43rd Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, 1–6.

Bradshaw, C., Atkinson, S., & Doody, O. (2017). Employing a qualitative description approach in health care research. *Global Qualitative Nursing Research*, *4*, 2333393617742282.

Bringer, M. L., Chelmecki, C. A., & Fujinoki, H. (2012). A survey: Recent advances and future trends in honeypot research. *International Journal of Computer Network and Information Security*, *4*(10), 63.

Chapman, J. (2019). *How Safe is Your Data?: Cyber-security in Higher Education*. Higher Education Policy Institute.

Chetalam, L. J. (2018). *Enhancing Security of MPesa Transactions by Use of Voice Biometrics*. United States International University-Africa.

Cohen, F., & Koike, D. (2003). Leading attackers through attack graphs with deceptions. *Computers & Security*, *22*(5), 402–411.

Conrad, E., Misenar, S., & Feldman, J. (2012). *CISSP study guide*. Newnes.

Cranford, E. A., Gonzalez, C., Aggarwal, P., Cooney, S., Tambe, M., & Lebiere, C. (2020). Toward Personalized Deceptive Signaling for Cyber Defense Using Cognitive Models. *Topics in Cognitive Science*, *12*(3), 992–1011.

de Bruijn, H., & Janssen, M. (2017). Building Cybersecurity Awareness: The need for evidence-based framing strategies. *Government Information Quarterly*, *34*(1), 1–7. https://doi.org/10.1016/j.giq.2017.02.007

Dlamini, M. T., Venter, H. S., Eloff, J. H. P., & Eloff, M. M. (2020). *Digital deception in cybersecurity: An information behaviour lens*.

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1–4.

Ferguson-Walter, K. J., LaFon, D. S., & Shade, T. B. (2017). Friend or faux: Deception for cyber defense. *Journal of Information Warfare*, *16*(2), 28–42.

Fielder, J. D. (2021). Cyber security in Kenya: Balancing economic security and internet freedom. In *Routledge Companion to Global Cyber-Security Strategy* (pp. 543–552). Routledge.

Friedberg, I., Skopik, F., Settanni, G., & Fiedler, R. (2015). Combating advanced persistent threats: From network event correlation to incident detection. *Computers and Security*, *48*, 35–57. https://doi.org/10.1016/j.cose.2014.09.006

Fugate, S., & Ferguson-Walter, K. (2019). Artificial intelligence and game theory models for defending critical networks with cyber deception. *AI Magazine*, *40*(1), 49–62.

González-Granadillo, G., González-Zarzosa, S., & Diaz, R. (2021). Security Information and Event Management (SIEM): Analysis, Trends, and Usage in Critical Infrastructures. *Sensors*, *21*(14), 4759.

Guarte, J. M., & Barrios, E. B. (2006). Estimation under purposive sampling. *Communications in Statistics-Simulation and Computation*, *35*(2), 277–284.

Hamed, H. H., El-Atawy, A., & Al-Shaer, E. (2006). Adaptive Statistical Optimization Techniques for Firewall Packet Filtering. *INFOCOM*, *6*, 1–12.

Heckman, K. E., Stech, F. J., Thomas, R. K., Schmoker, B., & Tsow, A. W. (2015). Cyber denial, deception and counter deception. *Advances in Information Security*.

Hutchins, E. M., Cloppert, M. J., & Amin, R. M. (2011). Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. *Leading Issues in Information Warfare & Security Research*, *1*(1), 80.

Jajodia, S., Cybenko, G., Liu, P., Wang, C., & Wellman, M. (2019). *Adversarial and Uncertain Reasoning for Adaptive Cyber Defense: Control-and Game-theoretic Approaches to Cyber Security* (Vol. 11830). Springer Nature.

Johansson, J. (2019). *Countermeasures Against Coordinated Cyber-Attacks Towards Power Grid Systems: A systematic literature study*.

Kashorda, M., & Waema, T. (2014). E-Readiness survey of Kenyan Universities (2013) report. *Nairobi: Kenya Education Network*.

Kaspersky. (2019). What is Cyber Security? | Definition, Types, and User Protection | Kaspersky. In *Www.Kaspersky.Co.Uk* (p. Home/Resource Centre/Definitions).

Khraisat, A., Gondal, I., Vamplew, P., & Kamruzzaman, J. (2019). Survey of intrusion detection systems: Techniques, datasets and challenges. *Cybersecurity*, *2*(1). https://doi.org/10.1186/s42400-019-0038-7

Klenka, M. (2021). Aviation cyber security: Legal aspects of cyber threats. *Journal of Transportation Security*, 1–19.

Maranga, M. J., & Nelson, M. (2019). Emerging Issues in Cyber Security for Institutions of Higher Education. *International Journal of Computer Science and Network*, *8*(4), 371–379.

Masai, J., & Wanja, M. (2016). *Securing Campus Wireless LANs*.

Min, B., & Varadharajan, V. (2016). A novel malware for subversion of self-protection in anti-virus. *Software: Practice and Experience*, *46*(3), 361–379.

Mirilla, D. F. (2018). *Slow Incident Response in Cyber Security: The Impact of Task Disengagement in Security Operations Centers*. Pace University.

Mwathi, D. G. (2018). *A model based approach for implementing authentication and access control in public WLANs: A case of Universities in Kenya*.

Nwanze, N., & Summerville, D. (2008). Detection of anomalous network packets using lightweight stateless payload inspection. *2008 33rd IEEE Conference on Local Computer Networks (LCN)*, 911–918.

Oyelaran-Oyeyinka, B., & Adeya, C. N. (2004). Internet access in Africa: Empirical evidence from Kenya and Nigeria. *Telematics and Informatics*, *21*(1), 67–81.

Paquet, C. (2012). *Implementing Cisco IOS Network Security (IINS 640-554) Foundation Learning Guide: Imp Cisco IOS Netw Sec F _c2*. Cisco Press.

Perkins, R. C., & Howell, C. J. (2021). Honeypots for Cybercrime Research. In *Researching Cybercrimes* (pp. 233–261). Springer.

Pfleeger, S. L., & Stolfo, S. J. (2009). Addressing the insider threat. *IEEE Security & Privacy*, *7*(6), 10–13.

Provos, N. (2004). A Virtual Honeypot Framework. *USENIX Security Symposium*, *173*(2004), 1–14.

Rowe, N. C., Custy, E. J., & Duong, B. T. (2007). Defending cyberspace with fake honeypots. *J. Comput.*, *2*(2), 25–36.

Serianu. (2020). *Africa Cybersecurity Report, Kenya 2019/2020*.

Stahnke, M. (2006). Legacy Protocols: Why Replace Telnet, FTP, rsh, rcp, and rlogin with SSH? *Pro OpenSSH*, 3–15.

Sun, J., Liu, S., & Sun, K. (2019). A scalable high fidelity decoy framework against sophisticated cyber atacks. *Proceedings of the ACM Conference on Computer and Communications Security*, 37–46. https://doi.org/10.1145/3338468.3356826

Sun, J., Sun, K., & Li, Q. (2017). CyberMoat: Camouflaging critical server infrastructures with large scale decoy farms. *2017 IEEE Conference on Communications and Network Security, CNS 2017*, *2017-Janua*, 1–9. https://doi.org/10.1109/CNS.2017.8228642

Symantec, C. (2017). *Internet security threat report: Volume 22*.

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, *48*(6), 1273–1296.

Tejada, J. J., & Punzalan, J. R. B. (2012). On the misuse of Slovin's formula. *The Philippine Statistician*, *61*(1), 129–136.

Thapa, S., & Mailewa, A. (2020). *The Role of Intrusion Detection/Prevention Systems in Modern Computer Networks: A Review*.

Trassare, S. T., Beverly, R., & Alderson, D. (2013). A technique for network topology deception. *MILCOM 2013-2013 IEEE Military Communications Conference*, 1795–1800.

Tyagi, M. (2017). *Security against cyber-crime: Prevention and detect*. Horizon Books (A Division of Ignited Minds Edutech P Ltd). https://books.google.co.ke/books?id=MMpJDwAAQBAJ

Updyke, D. D., Dobson, G. B., Podnar, T. G., Osterritter, L. J., Earl, B. L., & Cerini, A. D. (2018). *GHOSTS in the Machine: A Framework for Cyber-Warfare Exercise NPC Simulation*.

Vasilomanolakis, E., Karuppayah, S., Mühlhäuser, M., & Fischer, M. (2015). Taxonomy and survey of collaborative intrusion detection. *ACM Computing Surveys (CSUR)*, *47*(4), 1–33.

Vrable, M., Ma, J., Chen, J., Moore, D., Vandekieft, E., Snoeren, A. C., Voelker, G. M., & Savage, S. (2005). Scalability, fidelity, and containment in the Potemkin virtual honeyfarm. *Operating Systems Review (ACM)*, *39*(5), 148–162. https://doi.org/10.1145/1095809.1095825

Wang, C., & Lu, Z. (2018). Cyber deception: Overview and the road ahead. *IEEE Security & Privacy*, *16*(2), 80–85.

Whyte, C. (2020). Poison, Persistence, and Cascade Effects. *Strategic Studies Quarterly*, *14*(4), 18–46.

Zeinali, S. M. (2016). *Analysis of security information and event management (SIEM) evasion and detection methods*. Master Thesis, Tallinn University of Technology.

Zhang, X., Tsang, A., Yue, W. T., & Chau, M. (2015). The classification of hackers by knowledge exchange behaviors. *Information Systems Frontiers*, *17*(6), 1239–1251.

APPENDICES

Appendix 1: Data collection instruments

Questionnaire

UNIVERSITY OF EMBU

DEPARTMENT OF MATHEMATICS, COMPUTING AND INFORMATION TECHNOLOGY

MASTERS RESEARCH

QUESTIONNAIRE – PRELIMINARY SURVEY

Kindly respond to the following questions. The responses will be treated with the utmost confidentiality and will only be used to develop the theory for the research.

Part A: Demography

(1) Date……
……..

(2) University
Name……...

(3) Designation…….

Part B: University Cyber Security Awareness

(1) Do you have a Cyber Security and IT infrastructure in the university?

Yes	No

If YES to (1) above, estimate the number of systems that operate under the infrastructure……

(2) Indicate the number of IT personnel working specifically in cybersecurity……

(3) Kindly name any four university systems that are accessed online via the web……
……
……

(4) Are you aware of any security features employed on the university systems and networks?

Yes	No

(5) In your opinion, does the university have adequate cybersecurity policies in place?

Yes	No

Part C: University Cyber Security tools and techniques

(Tick the most appropriate option).

(1) Does the university have a cybersecurity program?

Yes	No

(2) If you already have a cybersecurity program in place, is it: (**Select all that apply**)

a) In house

b) Outsourced / Through a managed service

c) Don't know

d) Other (please specify)……..

(3) Which tools are used in network security and data protection in the university?

a) Access control

b) Anti-malware software

c) Anomaly detection

d) Application security

e) Data loss prevent (DLP)

f) Email security

g) Endpoint security

h) Firewall

i) Intrusion prevention systems

j) Intrusion Detection systems

k) Security information and event management (SIEM)

l) Virtual private network (VPN)

m) Web security

n) Wireless security

o) Other (please specify)……..

(4) From the selected tools in (3) above, do you understand how they operate?

Yes	No

(5) If YES, from your experience, how effective are they in preventing cyber-attacks?

a) Very useful

b) Not useful

c) I don't know

d) Other (please specify)…….

(6) Are all the cybersecurity tools open source?

Yes	No

(7) If the answer in (6) above is a NO, how did the university acquire the tools?

a) Procurement

b) Donation

c) Partnerships

d) Grants

e) Other (please specify)……..

(8) In your own opinion, do you think the tools used need to be improved?

Yes	No

Part D: University Cyber threats detection and analysis

(1) What are the common cyber threats affecting the university?

a) Phishing

b) Malware

c) Denial of service

d) Man-in-the-middle attack

e) Other (please specify)……..

(2) How do the cybersecurity systems in place detect the attacks in the university?

Through;

a) Firewall

b) IDS

c) IPS

d) SIEMs

e) Decoys

f) Honeypots

g) None

h) Other (please specify)……

(3) How many cyber-attack cases have you received in the past 12 months?

Yes	No

a) Below 100
b) Between 100 and 500
c) Over 500
d) None

(4) How does the cybersecurity team respond to attacks when notified? Explain

……
……
……
……
……
……
……
……
……

(5) Which systems are highly targeted by cybercriminals?

a) Enterprise resource planning
b) Service portals
c) Payroll management system
d) Library management system
e) Other (please specify)…….

(6) Does the response team keep a database of all the attacks?

Yes	No

(7) Suppose the answer in (9) above is yes. Explain how they make use of the database.

.....
.....
.....

.....

.....

(8) Please indicate which threat management platform(s) you use, if any. (**Select all that apply**)

a) Identity and Access Management (IAM)
b) Endpoint security
c) CTI service provider
d) Deception-based detection
e) SIEM
f) Network packet broker/ Inline monitoring
g) IDS/IPS/UTM/ firewall
h) None

(9) How do the cyber-attacks and threats being analyzed in the university? Using

a) Online software tools
b) Customized application
c) Not analyzed
d) Other (please indicate)……

(10) Do you understand how deception decoy

Yes	No

work?

(11) If the answer in (3) is a YES, explain how you think it can be incorporated in the university network and systems to improve threat management and monitoring of attacks.

……

……

……

Thank you for taking the time to respond

Appendix 2:Authorization letter from BPS

UNIVERSITY OF EMBU

OFFICE OF THE DIRECTOR
BOARD OF POSTGRADUATE STUDIES

Tel. 0727933950, 0788199505
Website: www.embuni.ac.ke

P.O. Box 6-60100, Embu
E-mail: dir.bps@embuni.ac.ke

Our Ref: B529/1158/2017
Your Ref:

Date: 10th September 2020

Serem, Kiprono Edwin,
P/a
Department of Mathematics, Computing and Information Technology,

Dear Mr. Serem.

RE: APPROVAL OF RESEARCH PROPOSAL

This is to inform you that the Board of Postgraduate Studies, at its meeting of 21st July 2020, approved your research proposal for M.Sc Degree entitled "Protecting Institutions of Higher Learning in Kenya: A Scalable Hybrid Decoy Framework Against Cyber threats" You may now proceed with your data collection subject to obtaining a research permit from NACOSTI.

As you embark on your data collection, please note that you are required to:

i. Consult your supervisor(s) at least once a month.
ii. Submit to the Board of Postgraduate Studies at least two (2) duly completed Postgraduate Progress Report Forms through the Chairman of Department and Dean of School every three (3) months.
iii. Give a minimum of two (2) seminar presentations before submission of thesis.
iv. Publish at least one (1) paper before the project report/thesis is submitted for examination.
v. Adhere to the University Plagiarism Policy and the prescribed similarity levels.
vi. Obtain other permits, permission or clearance such as ERC, IBC, KWS if required.

The Progress Report Forms, research project/thesis submission checklist and other important postgraduate documents are available at the University's website under Board of Postgraduate Studies webpage http://bps.embuni.ac.ke/ as downloads.

Thank you.

Prof. Nancy Budambula
DIRECTOR, BOARD OF POSTGRADUATE STUDIES
NB/dk

Copies to:
1. DVC (ARE)
2. Registrar, ARE
3. Dean, SPAS
4. CoD,MCIT
5. Supervisors:Dr.David M.Mugo& Dr.Boaz Kipyego Too

Appendix 3:Letter from NACOSTI

REPUBLIC OF KENYA

NATIONAL COMMISSION FOR SCIENCE,TECHNOLOGY & INNOVATION

Ref No: 265982

Date of Issue: 30/September/2020

RESEARCH LICENSE

This is to Certify that Mr.. Soren Kiprono Edwin of University of Embu, has been licensed to conduct research in Bomet, Bungoma, Busia, Embu, Garissa, Homabay, Kakamega, Kericho, Kiambu, Kilifi, Kirinyaga, Kisii, Kisumu, Kitui, Kwale, Laikipia, Lamu, Machakos, Makueni, Meru, Mombasa, Muranga, Nairobi, Nakuru, Nandi, Narok, Nyamira, Nyandarua, Nyeri, Taita-Taveta, Tharaka-Nithi, Turkana, Uasin-Gishu, Vihiga on the topic: PROTECTING INSTITUTIONS OF HIGHER LEARNING IN KENYA: A SCALABLE HYBRID DECOY FRAMEWORK AGAINST CYBER THREATS for the period ending : 30/September/2021.

License No: NACOSTI/P/20/6961

265982

Applicant Identification Number

Director General
NATIONAL COMMISSION FOR SCIENCE,TECHNOLOGY & INNOVATION

Verification QR Code

NOTE: This is a computer generated License. To verify the authenticity of this document, Scan the QR Code using QR scanner application.

Appendix 4:List of all public/private Universities

1. University of Nairobi
2. Kenyatta University
3. Strathmore University
4. Jomo Kenyatta University of Agriculture and Technology
5. United States International University Africa
6. Moi University
7. Egerton University
8. Mount Kenya University
9. The Catholic University of Eastern Africa
10. Murang'a University of Technology
11. Kenya Methodist University
12. Africa Nazarene University
13. Technical University of Kenya
14. Maseno University
15. Daystar University
16. University of Eastern Africa,
17. Kabarak University
18. Machakos University
19. Multimedia University of Kenya
20. Dedan Kimathi University of Technology
21. Masinde Muliro University of Science and Technology
22. KCA University
23. St. Paul's University
24. Riara University
25. South Eastern Kenya University
26. Chuka University
27. Maasai Mara University
28. Pwani University
29. University of Eldoret
30. Kisii University
31. Africa International University
32. Jaramogi Oginga Odinga University of Science and Technology

33. Karatina University
34. University of Embu
35. Kibabii University
36. Technical University of Mombasa
37. The Co-operative University of Kenya
38. International Leadership University, Kenya
39. University of Kabianga
40. Pan Africa Christian University
41. Meru University of Science and Technology
42. Management University of Africa
43. Garissa University
44. The Presbyterian University of East Africa
45. Amref International University
46. Adventist University of Africa
47. Pioneer International University
48. Zetech University
49. Laikipia University
50. KAG East University
51. Great Lakes University of Kisumu
52. Kiriri Women's University of Science and Technology
53. Umma University
54. Kirinyaga University
55. Rongo University
56. Scott Christian University
57. Taita Taveta University
58. Gretsa University
59. Kenya Highlands University
60. Lukenya University
61. The East African University
62. RAF International University

Appendix 5: Contribution of the Study

Cyber security is a vast field and requires the input of many stakeholders. The technical team are the combat personnel, while the users are prone to attacks. This study has addressed the challenges encountered by the cyber defence teams by enhancing deceptive decoys techniques to further make the system complex for the attackers. In addition, the survey done in the universities and colleges gives the policymakers basis of more emphasis on cyber security in institutions of higher learning. The automatic generation of user activities and networks will build knowledge of how decoys can apply deceptive mechanisms. Finally, the use of open-source resources has been identified as easier for organisations that, for one reason or the other, may not have the financial capacity to purchase the expensive proprietary cyber defence tools.

Appendix 6: publication

International Journal of Electrical Engineering and Technology (IJEET)
Volume 12, Issue 6, June 2021, pp. 281-292, Article ID: IJEET_12_06_027
Available online at https://iaeme.com/Home/issue/IJEET?Volume=12&Issue=6
ISSN Print: 0976-6545 and ISSN Online: 0976-6553
DOI: 10.34218/IJEET.12.6.2021.027

© IAEME Publication Scopus Indexed

DECEPTIVE DECOYS: COMBINING BELIEVABLE USER AND NETWORK ACTIVITIES AND DECEPTIVE NETWORK SETUP IN ENHANCING EFFECTIVENESS

Edwin K. Serem
Department of Mathematics, Computing and Information Technology, University of Embu, 60100 Embu, Kenya

David M. Mugo
Department of Mathematics, Computing and Information Technology, University of Embu, 60100 Embu, Kenya

Boaz K. Too
Department of Mathematics, Computing and Information Technology, University of Embu, 60100 Embu, Kenya

ABSTRACT

Cybersecurity threats are a malicious act that seeks to damage, steal, or gain unauthorized access to information. In recent years there has been an attempt by cybersecurity specialists to come up with an effective system that proactively protects the systems from cyber-attacks. Cyber deception is one efficient method that makes use of decoys to entrap attacks and divert them from real systems. However, existing cyber decoys lack efficiency in hiding true identity due to impractical user activity and network simulation. In this paper, we propose a hybrid decoy system that combines the use of two-layered decoys in the front-end and back-end with an SSH tunnel in between. The front-end decoys will capture attacks and forward them to backend decoys for execution and feedback. General HOSTS framework was used to generate believable user and network activities that can effectively convince the attackers that they are attacking the real systems. All attacker activities are logged by Logstash and presented using Grafana with the Kibana user interface. The experimental results demonstrate that our system can effectively misdirect and misinform attackers by combining deceptive network setup and configurations as well as generating fake user and network activities.

Key words: Decoy, Honeypot, advanced persistent threats (APT), Virtual Machine (VM), Linux Containers (LXC)