Media Engineering and Technology Faculty
German University in Cairo

Comparative Study of Anomaly Detection and Signature-Based Intrusion Detection Systems

Bachelor Thesis

Author:	Sara Mohamed
Supervisors:	Dr Sherif Saad
Submission Date:	18 May 2024

Media Engineering and Technology Faculty
German University in Cairo

Comparative Study of Anomaly Detection and Signature-Based Intrusion Detection Systems

Bachelor Thesis

Author: Sara Mohamed
Supervisors: Dr Sherif Saad
Submission Date: 18 May 2024

This is to certify that:

(i) the thesis comprises only my original work toward the Bachelor's Degree

(ii) due acknowledgment has been made in the text to all other material used

Sara Mohamed
18 May 2024

Acknowledgments

First and foremost, I would like to express my gratitude to Dr. Sherif Saad for allowing me to work on this project, as well as for his efforts and consistent follow-up and guidance throughout the semester. And for sure, I would like to thank my parents for their patience, attention, and their support for allowing me to be here today. Also, I will not forget my friends for their motivation and support. Without all of these people, nothing would have been done.

Abstract

Cyberattacks are on the rise, posing increasingly daunting challenges in the accurate detection of intrusions. Failing to prevent these intrusions can degrade the credibility of security services, such as safeguarding data confidentiality, integrity, and availability.

The role of an Intrusion Detection System (IDS) is to continuously monitor our network for signs of malicious activity and deviations from normal behavior. Our objective extends beyond merely discussing cybersecurity data science and relevant methodologies; we aim to emphasize the practical application of data-driven intelligent decision-making to safeguard systems against cyber threats.

In this study, we delve into two primary types of IDS: Anomaly-based Detection Systems (ANIDS) and Signature-based Detection Systems. To distinguish between these systems, Signature-based detection relies on identifying known threats using a predefined list, while Anomaly-based detection analyzes normal system behavior to establish a baseline. When deviations occur, it alerts security personnel to potential threats.

Contents

Chapter 1

Introduction

1.1 Background

As technology progresses and our reliance on the internet grows, we encounter a significant challenge: network attacks. Nowadays, people heavily depend on the internet for communication, which exposes confidential data exchanged between parties to potential theft or corruption. Network attacks take various forms, prompting enterprises to maintain powerful cybersecurity standards, enforce robust network security policies, and provide comprehensive staff training to safeguard their assets against increasingly sophisticated cyber threats.

1.1.1 Network Attacks

A network attack refers to the malicious exploitation of any vulnerabilities present within a network, with the intent to steal, damage, or gain unauthorized access to the system. These attacks typically fall into two primary categories: Passive Attacks and Active Attacks. Passive Attacks involve intruders gaining unauthorized access to data and simply monitoring or eavesdropping on a system or network. Conversely, Active Attacks involve intruders altering or deleting stolen data. Various forms of Active Attacks include:

1. Session Hijacking Attack
2. Message Modification Attack
3. Masquerade Attack
4. Denial-of-Service Attack
5. Distributed Denial-of-Service Attack
6. Trojans

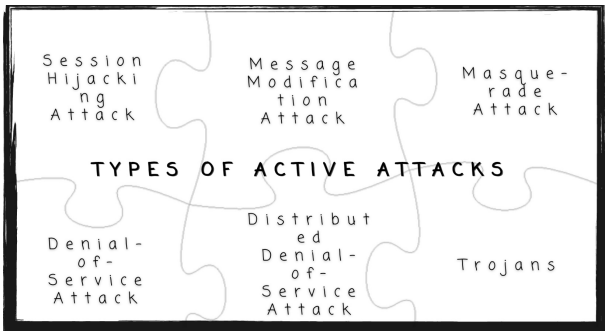

Figure 1.1: Types of Active Attacks
https://www.zenarmor.com/docs/network-security-tutorials/what-is-active-attack

We'll begin by addressing the Session Hijacking Attack, also termed cookie hijacking or cookie side-jacking, classified as an active attack. In this attack, unauthorized individuals gain entry to a user's session ID, enabling them to effectively impersonate the user. Usually, attackers exploit session cookies employed in the HTTP communication protocol to identify browsers. These session cookies remain stored in the browser until the user manually logs out or is automatically logged out. Hackers endeavor to access these session cookies to execute a session hijacking attack.

Secondly, We have a Message Modification Attack which is an attack where the intruder alters the data sent between two parties this is done by the attacker snooping in on the communication channel between the two parties and then changing the data sent or adding fake data.

Thirdly, we have a Masquerade Attack which is defined by its name where the attacker pretends to be a legitimate user in order to log in to the system. This could be done by Phishing emails where the intruder sends fake emails for the user to fall into the trap of revealing their secured credentials.

The fourth type is the Denial-of-Service Attack, often abbreviated as DOS, which aims to disrupt the normal functioning of a particular server, service, or network. This is accomplished by flooding the targeted server with an excessive volume of traffic, ultimately causing it to crash.

Fifthly, there is the Distributed Denial-of-Service Attack, commonly referred to as DDOS, which shares similarities with the Denial-of-Service Attack. The key distinction lies in the fact that in a DDOS attack, multiple devices situated across different locations collectively target the desired server, as opposed to a single device in a traditional DOS attack.

Lastly, there's the Trojan, much like the sneaky ploy in Greek mythology where adversaries snuck into Troy by hiding inside a giant wooden horse offered as a gift. Just like that tactic, this virus, taken after the Trojan horse, slips into systems through innocent-looking email attachments, essentially barging in and snooping on user data and activities, much like unwanted guests rummaging through your belongings.

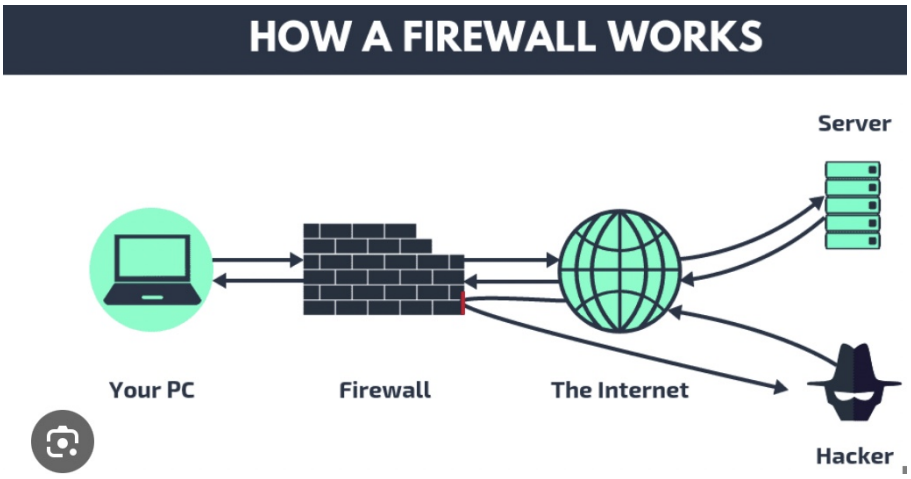

Figure 1.2: How a Firewall Works
https://www.linkedin.com/pulse/how-does-firewall-work-ashwin-harish-p

1.1.2 CyberSecurity

The Need for Cybersecurity

Having identified various cyber threats, it's crucial to secure our systems and devices. This realm of protection is known as cybersecurity. It acts as a shield, defending computers, servers, programs, and networks from unauthorized access and malicious attacks.

Strategies for a Strong Cybersecurity Posture

There are several strategies to strengthen cybersecurity. Enforcing strong passwords and multi-factor authentication (MFA) makes hacking attempts more difficult. Also, regularly backing up data ensures recovery in case of a breach.

Firewalls, Encryption, and Intrusion Detection

Firewalls form a primary defense by filtering traffic entering and leaving your network. They can block malicious emails, malware downloads, and access to risky sites. Firewalls come in two forms: hardware versions built into routers for easy network protection and software versions installed on individual devices.

Another crucial technique is data encryption, which scrambles readable data into an unreadable format. This requires a cryptographic key, shared between sender and recipient, for decryption. The more complex the key, the stronger the encryption.

1.1.3 Intrusion Detection System (IDS)

Intrusion detection systems (IDS) function as critical security tools, continuously monitoring network traffic for signs of malicious activity or policy breaches. When an IDS detects suspicious behavior patterns, it triggers immediate alerts, notifying IT security

personnel for further investigation and response. Unlike firewalls that rely on predefined rules to block known threats, IDSs excel at identifying zero-day attacks and novel malware variants by analyzing traffic anomalies and system behaviors.

IDS technology employs various techniques for threat detection, including:

- **Signature-based detection:** This method compares network traffic patterns against a database of known malicious signatures. It's effective against well-documented threats but may miss zero-day attacks.
- **Anomaly-based detection:** This approach analyzes traffic for deviations from normal network activity patterns. It can identify novel threats but might generate false positives due to unusual but legitimate network usage.
- **Stateful protocol analysis:** This technique examines traffic for adherence to established network protocols. Deviations from protocol expectations can indicate potential attacks.

There are two primary categories of IDSs, each serving a distinct purpose:

- **Network Intrusion Detection Systems (NIDS):** These systems act as network guardians, strategically positioned at network perimeters (often behind firewalls) to monitor all incoming and outgoing traffic across the network. NIDS can detect attempts to breach network defenses by analyzing traffic for malicious content or suspicious behavior. Additionally, NIDS can be deployed internally to monitor for threats originating from compromised user accounts or insider actions.
- **Host Intrusion Detection Systems (HIDS):** Unlike NIDS which focus on network-wide traffic, HIDS are deployed on individual endpoints such as servers, laptops, and routers. They monitor system activity, file changes, and incoming/outgoing traffic on those specific devices, providing in-depth insights into potential threats targeting individual devices.

By deploying a combination of NIDS and HIDS, organizations can establish a comprehensive security posture, safeguarding against a wider range of threats – both external attacks attempting to infiltrate the network and internal threats targeting specific devices.

There are three main variant types of IDS which are Anomaly-based IDS (AIDS) , Signature-based IDS (SIDS) and Hybrid Intrusion Detection Systems (HIDS).

1.1.4 Anomaly-based IDS (AIDS)

It is an intrusion Detection System in which it forms a baseline of normal behavior for a network and then analyzes the system performance in comparison with the baseline. if any deviations appear an alert will occur notifying security administrators. The greatest benefit of the Anomaly-based IDS is that it can detect Zero-Day attacks while Signature-based IDS can't do that. Zero-Day attacks are cyber attack vector that preys on an undiscovered or unfixed security hole in computer hardware, firmware, or software

	Advantages	Drawbacks	Data sources	Examples
Host	Ability to detect malicious or improper activities of all the users	They can get compromised as soon as the host server is compromised by an attack	Audits records	PortSentry[13]
	Very effective in detecting suspicious user behaviors	Cannot monitor network traffic	Security audit information	
	Ability to operate in encrypted environments	High system resources consumption	System log files	Lhotsky[14]
		High setting-up and management costs	System accounting	
		Portability issue	System calls	Kim et al[15]
			System configuration	
Network	Extremely portable	Scalability issue	Simple network management protocol	RealSecure[16]
	Independent of the installed operating systems	Cannot operate in encrypted environments	Network connections	SecureNet[17]
	Real-time detection		Network traffic packets	Roesch et al[18]
	Low deployment and management costs			

Figure 1.3: Host VS Network
https://www.intechopen.com/chapters/8695

1.1.5 Signature-based IDS (SIDS)

It is a type of Intrusion Detection System where it acts by using pattern-matching methods to identify a recognized attack. In other words, when an intrusion signature matches the signature of a previous intrusion that already exists in the signature database, an alarm signal is triggered. SIDS typically demonstrates high accuracy in detecting intrusions that are already known. However, it struggles to identify zero-day attacks because there is no existing signature in the database until the signature of the new attack is identified and added.

1.1.6 Hybrid IDS

It is a more advanced intrusion Detection System that combines Anomaly-based IDS and Signature-based IDS called. The crucial function of Hybrid Intrusion Detection Systems (HIDS) that makes it different compared to both Anomaly-based IDS and Signature-based IDS is that it overcomes the limitations of each individual IDS making it more suitable nowadays.

1.2 Motivation

The main purpose of this paper is to review and compare the two intrusion detection systems and to make us understand the strengths and weaknesses of each approach, ultimately leading to more effective intrusion detection overall. One of the benefits of this comparative study is knowing the strengths and weaknesses of both approaches and then

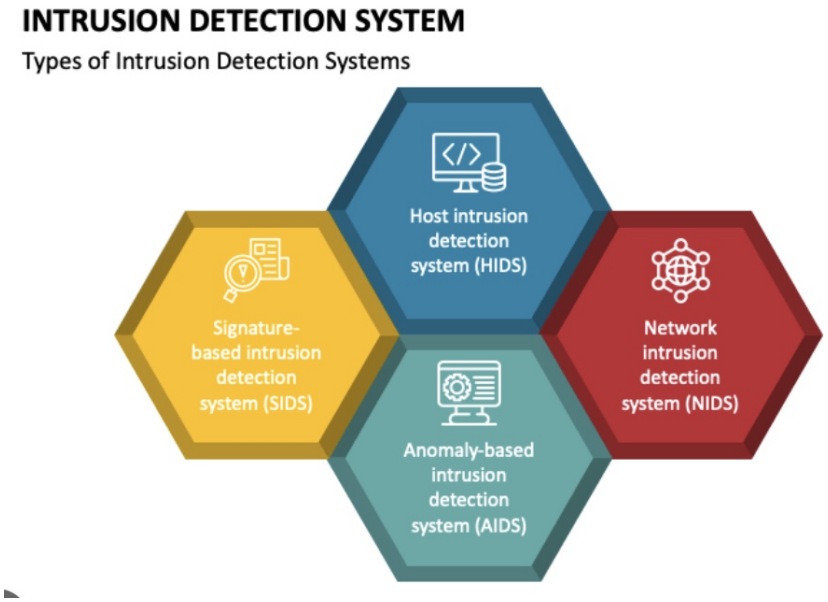

Figure 1.4: Types of Intrusion Detection Systems

$https://sdcsalesar.live/product_details/2101663.html$

organizations can make a definite decision of which approach to utilize that will suit them best. Another motive of this study is it allows benchmarking the performance of both anomaly-based intrusion detection systems and signature-based intrusion detection systems and identifies areas for improvement in both of them.

There are numerous factors that contribute to the need for funding these studies. One of them is that knowing and understanding both IDS will lead to organizations knowing which IDS is perfect for their needs and not losing time and money deploying each one. Also, funding such papers will encourage the development of new Intrusion Detection systems that solve the weaknesses of the old ones and add new features to the old ones that will solve present issues.

1.3 Expected Outcome

In the Comparative Analysis Report, We will have a comprehensive overview about the advantages and disadvantages of both Anomaly- based Intrusion Detection System and Signature-based Intrusion Detection System and also compare between both of them using multiple different criteria.

In the Recommendations, I will offer guidance to businesses or organizations on determining the most suitable type of IDS that aligns with their needed requirements.

In the Future Directions, We will provide insights into future trends and potential developments in the field of intrusion detection systems.

1.4 Methodology

In the Literature Review, I will provide a comprehensive review of existing researches, case studies, and industry reports on Anomaly-based Intrusion Detection System and Signature-based Intrusion Detection System.

In the System Selection, I will choose representative systems or tools for each type of IDS for in-depth analysis and comparison.

In the Criteria Development, I will develop a set of criteria to compare between the IDS types, which could include detection accuracy, speed, adaptability to new threats, false positive rates, resource requirements, and ease of implementation.

In the Empirical Analysis, We will set up a controlled environment to test and compare the selected IDS systems based on the developed criteria.

1.5 Thesis Structure

In Chapter 1, an introductory exploration of Intrusion Detection Systems (IDS) will be presented, emphasizing their indispensable role in cybersecurity infrastructure. The necessity for such systems will be underscored, outlining the critical need for proactive measures against evolving cyber threats. Additionally, an in-depth examination of the three primary types of IDS will be conducted. Furthermore, the chapter will discuss the motivation behind the research and the expected outcomes.

In Chapter 2, we will preview previous works related to Anomaly-based intrusion detection systems, Signature-based intrusion detection systems, and IDS. We will discuss the related works extracting the models used, the paper claim, a summary of the paper used, and the result of each paper.

In Chapter 3, a comprehensive summary of the previous literature review will be presented, followed by an in-depth exploration of the three primary types within each intrusion detection system category. Additionally, this chapter will detail the experimental methodology involving the utilization of Snort and Zeek across diverse datasets. The experiment will be meticulously structured to evaluate key criteria such as detection accuracy, speed, and ease of implementation, providing valuable insights into the efficacy of these intrusion detection systems.

Chapter 4, I will offer the conclusive findings of the experiments conducted, accompanied by an insightful discussion on the encountered challenges. Furthermore, it will outline prospective avenues for future research endeavors within the field, providing valuable insights for researchers aiming to build upon this study.

Chapter 2

Related Work

In this section, we will preview previous works related to Anomaly-based intrusion detection systems, Signature-based intrusion detection systems, and IDS. We will discuss the related works extracting the models used, the paper claim, a summary of the paper used, and the result of each paper.

2.1 Intrusion Detection System

- In [10] This paper presents a review of the research trends in network-based intrusion detection systems (NIDS), their approaches, and the most common datasets used to evaluate IDS Models. The analysis presented in this paper is based on articles published with their citations and most cited research articles related to the intrusion detection system in journals and conferences. Based on the published articles in the intrusion detection field for the last 15 years, this article also discusses the state-of-the-art NIDS, commonly used NIDS, citation-based analysis of benchmark datasets, and NIDS techniques used for intrusion detection. Also, a comparative analysis is presented in this paper to quantify the popularity of various approaches. Finally, this study is helpful for researchers interested in evaluating research trends in NIDS and their related applications.

  This paper focuses on research trends in previous articles (from 2005 to 2020) in the field of IDS, its related techniques, datasets, total publications, and other citation-related analyses. Also, the paper claims that an intrusion detection system (IDS) is a strong network security system capable of detecting unauthorized and abnormal network traffic flow, therefore enhancing network security.

  To prove that an intrusion detection system (IDS) is a strong network security system capable of detecting unauthorized and abnormal activities. The article has a comparison of several renowned IDS (Such as Snort, Base, Suricata, and Bro) that are being used nowadays. Added to that, an analysis was made about known datasets with their advantages, disadvantages, and the total number of citations

from the year 2005 to 2020. Furthermore, the article discusses various performance metrics used to evaluate an intrusion detection system. These performance metrics are False Positive Rate (FPR), precision, and F-score. Also, There is an analysis done in the article about the most cited published articles.

In this paper, We explored a comprehensive and straightforward analysis for anyone who wants to compare various approaches used to design Network Intrusion Detection models. This review is established based on numerous research papers in different journals/publications between 2005 and 2020. In this article, we took citation as a quantitative measure to review the popularity of the intrusion detection system among various approaches. Added to that, This paper presents various tables that offer a rapid analysis of different NIDS, research trends, and research scope. A review of diverse datasets with their characteristics, merits, demerits, and citation analysis has also been presented. The various approaches used in the network intrusion detection system are tabulated with their advantages and disadvantages and A review concerning research trends regarding different techniques in IDS is also discussed. The comparative research trend analysis regarding intrusion detection systems for a network is also graphically presented based on citations and several published article counts. In my opinion, the paper appears to be strong in terms of its thorough analysis, quantitative approach, methodological rigor, comparative insights, and conclusions. It provides valuable information for researchers and practitioners interested in network-based intrusion detection systems. In my opinion, this paper is strong as it gives a detailed analysis of the intrusion detection system with previous studies.

- In [7] This paper focuses on how to improve internet environment security against powerful assaults like DoS (Denial of Service Attack)and cyber attacks. Denial of Service Attacks are a crucial unresolved issue in Internet networks as they can disrupt two or more authorized entities trying to communicate with each other. The entire focus of this research is on examining the approaches used to develop intrusion detection systems. It emphasizes the need for strong processing power and the challenges of managing computational costs in IDS operations. The document highlights the two main modes of IDS - anomaly-based and signature-based - and the importance of combining these techniques for increased protection against cyberattacks. Additionally, it mentions the use of machine learning in IDS for training and identifying intrusions. Various studies and research work on network intrusion detection systems are referenced, showcasing different methodologies and techniques employed by researchers. These include the use of Convolution Neural Networks (CNN), Decision Trees, Deep Learning, and other machine learning methods for detecting and mitigating intrusions. The document also touches upon the importance of feature selection algorithms, signature matching, prevention systems, threat reporting, anomaly detection, and different types of Intrusion Prevention Systems (IPS) in enhancing network security.

  The paper claims that enhancing protection for information and networks using Intrusion Detection Systems (IDS) is crucial in today's digital age where cyber

threats are consistent. Added to that, the article suggests that IDs combining both Anomaly-based and signature-based Techniques can help us detect intrusions in networks effectively. Also, the paper discusses the importance of using machine learning algorithms and training models with diverse datasets to improve the accuracy and efficiency of intrusion detection systems. Additionally, the paper suggests that a dynamic approach, incorporating various methodologies and techniques, is important for robust network security.

The paper gives a detailed analysis of various methods and techniques used by researchers to detect intrusions in information and network systems. These methods are anomaly, signature detection, and signature comparison. Also, in the related work section, we have multiple references that discuss different approaches such as Convolutional Neural Networks (CNN), Decision Trees, Deep Learning, and other machine learning methods for intrusion detection. It also discusses the use of feature selection algorithms, signature matching, prevention systems, threat reporting, anomaly detection, and different types of Intrusion Prevention Systems (IPS) to enhance network security. Furthermore, the paper compares and contrasts the effectiveness of different intrusion detection techniques based on the datasets used and the outcomes achieved. Added to that, It emphasizes the importance of merging multiple techniques and datasets for dynamic handling of intrusions, suggesting that specific methods should be applied based on the specific needs of the situation.

This paper gives a detailed analysis of methods that have been used by researchers to reduce network attacks. Researchers proposed different methods and techniques for handling attacks such as AI approaches and Prevention Systems. Also, there was an analysis of previous research where it opened our eyes that the future intrusion detection systems must be adaptable to various datasets and capable of integrating multiple techniques for dynamic handling of intrusions. Furthermore, we can conclude from this paper that it is not mandatory to utilize all methods simultaneously; instead, specific methods can be utilized based on the demands or requirements of the organization. In my opinion, the paper appears to be strong in terms of its coverage of various methodologies and techniques used in enhancing protection for information and networks using Intrusion Detection Systems (IDS).

- In [11] This paper analyses Intrusion Detection Systems nowadays. Therefore, it focuses on the IDS classification, which includes three types of classification which are Network-Based IDS type, Host-Based IDS type, and cloud- based solutions. Also, It discusses different categories of IDS such as Network-based IDS (NIDS), Host-based IDS (HIDS), Hybrid IDS, and Cloud IDS (CIDS). The paper focuses on ways in which we can detect intrusions and anomalies, providing examples of the best detecting tools (OSSEC, Snort, or Bro (ZEEK)). Another aspect covered in this paper is the comparison between three systems: IDS, IPS, and IRS, and their response. The paper also compares and contrasts NIDS, HIDS, and CIDS, highlighting their strengths and weaknesses in detecting and responding to intrusions and attacks. Finally, Intrusion Detection Systems' evasion techniques and challenges will be discussed followed by conclusions.

This paper's purpose is to educate people about various aspects of Intrusion Detection Systems (IDS) such as different categories of IDS, attack detection techniques, response strategies, and the importance of IDS in network security. Also, the paper focuses on the different three types of IDS (Host-based, Network-based, and Cloud-based) giving the advantages and disadvantages of each type. Added to that, the paper discusses the evasion techniques that attackers use to avoid being detected by security personnel. Also, it goes into detail about the meaning and the difference between IDS, IPS, and IRS (Intrusion Response System).

To educate people about Intrusion Detection Systems (IDS), the paper discusses the different types of IDSs like Network-based, Host-based, and Cloud-based in detail. Added to that, it discusses different ways to detect intrusions Anomaly Detection, Signature Detection, Heuristic Detection, and Hybrid Detection. Furthermore, The paper explores the difference between IDS, IPS, and IRS and then chooses which system of is the best to detect all intrusions and malicious attacks. Also, a comparison is done between three commonly used IDS tools (OSSEC, SNORT, and BRO) with different features. Finally, the paper focuses on defining various Evasion techniques such as (Encryption and Fragmentation) which are ways that attackers avoid being detected.

This paper explains Intrusion Detection Systems, their types, and detection techniques. Besides HIDS and NIDS types of IDS, there was a cloud method analyzed, which offers future solutions. Furthermore, the paper elaborated and compared IDS, IPS, and IRS in terms of their roles in the network and responses to the threat. Finally, the paper listed the most common IDS tools: OSSEC, Snort, and Bro, and compared each other according to specific features. This report's aim is not only to present facts about Intrusion Detection Systems but also to aid in understanding the importance of network protection and that there is not one correct solution that would provide security. In the paper, we analyzed HIDS, NIDS, and CIDS in terms of their way of work, merits, and demerits. Similarly, there were discovered: AIDS, SIDS, Heuristic-based, and Hybrid- based detections. Each of them standalone is not good enough, with a big room for development. However, using hybrid solutions supported by Machine Learning, where there could be combined two or three different systems as proposed before: IDPS or IDPRS, a hybrid of SIDS and AIDS, would create a properly built security layer. Certainly, each establishment lists different requirements and needs different security techniques/levels. However, they need to be aware of the rapid technology development and attackers' new ways to be detected – evasion techniques. These kinds of security challenges force us to develop new security technologies, reuse and improve the old ones, and look for completely new solutions like the proposed CIDS. In my opinion, this paper is strong as it discusses Intrusion Detection System categories, Attack Detection, and response in detail. It also gives a detailed overview of the best IDS tools with a comparison table and discusses the Evasion Techniques in detail.

- In [13] This paper discusses intrusion detection technologies, methodologies, and approaches and also investigates new attack types, protection mechanisms, and re-

cent scientific studies that have been made in this area. In addition, the paper focuses on available datasets, well-known IDS tools, and the advantages and disadvantages of particular IDSs in detail. Finally, this scientific review study presents a road map for researchers and industry employees who focus on IDSs.

This study aims to analyze IDSs, current development methods, available datasets, and remaining problems in detail. For this purpose, intrusion detection technologies, methodologies, commonly used tools, and leading methods are examined thoroughly in the literature review section. This paper also presents a detailed literature review to investigate and examine the current state of IDSs.

First of all, information about what this system is, and in general, the basic features that should be in an IDS are mentioned. Afterward, IDSs are categorized according to the way they monitor the network traffic, record flow data, detect attacks, and report warnings. All IDS technologies, methodologies, and approaches within this scope have been examined in detail. Their strengths and weaknesses are mentioned, and a comprehensive summary of the work done in each area is given. Then, the datasets that are widely used in the testing and evaluation phase of the developed intrusion detection systems were examined and detailed information was given about these datasets. Finally, common intrusion detection tools used by individuals, institutions, and organizations to recognize attacks are mentioned. The intrusion detection method used by each of the intrusion detection tools, their advantages and disadvantages are reviewed.

The paper discusses the most commonly used datasets in IDSs like (the KDD'99 Dataset). Also, intrusion detection systems are summarized based on the techniques that are used, different detection methods, and the main idea of each

detection approach. Then, available datasets, advantages, and disadvantages of each detection technology, and current state-of-the-art studies are analyzed. Finally, a comparison of the detection techniques and the future of the research directions, and our thoughts about IDSs are given.

The results of the paper discuss various aspects of Intrusion Detection Systems (IDSs), such as intrusion detection methodologies, different detection methods such as anomaly-based and network-based, wireless intrusion detection methods, stateful protocol analysis, and network-based intrusion detection methods.

Due to the comparison of IDSs and stating their advantages and disadvantages, the paper highlighted the need for improvement and integration of new technologies like cloud, machine learning, and deep learning. Added to that, the paper discusses existing challenges and proposes assumptions for enhancing IDSs, along with future research directions in the field of intrusion detection systems. Also, a detailed overview of IDSs will serve as a roadmap for researchers and industry professionals focusing on IDSs.

This review paper is different from the previous survey papers in many aspects. Previous studies are mainly focused on only one or two subjects such as intrusion detection methodologies or datasets that have been used. However, in this study,

the various aspects of the IDSs are discussed. Besides, several suggestions are being made for each subject. The paper also makes contributions not only to researchers but also to private companies that want to utilize IDSs more effectively.

In my opinion, the paper is strong according to its claim. The paper aims to present a comprehensive systematic literature review on Intrusion Detection Systems (IDSs).

- In [1]Network Intrusion Detection in Cloud, this paper discusses several approaches to Network Intrusion Detection in Cloud Environments. A comparison was made between these various approaches to determine the appropriate choice in regard to organizations' needs and their advantages and disadvantages. The findings of these studies demonstrate the effectiveness of conventional methods in known threat detection. Also, these findings aided us in understanding the need for hybrid approaches that integrate the strengths of both. Added to that, The paper addressed multiple challenges such as privacy compliance, performance scalability, and false positives. It also highlighted the importance of continuous monitoring, privacy-preserving technologies, and real-time threat intelligence integration. This study also highlights the importance of staff training for the successful implementation of a network intrusion detection system (NIDS). Furthermore, the paper discusses the future scope of NIDS in cloud computing environments to meet emerging attacks and regulatory requirements.

  The paper aims to gain a detailed understanding of the advantages and disadvantages of NIDS approaches like the Anomaly approach, AI approach, and behavior approach. It also gives a comprehensive overview of anomaly-based, signature-based, and behavior-based NIDS approaches in the cloud-based system. Added to that, The article presents a detailed analysis of various customized NIDS approaches, especially for the cloud environment. Furthermore, organizations carefully examine emerging trends and predict the potential future direction of threats in cloud systems. This article also aims to provide a practical direction for applying and optimizing this system within the cloud environment. Furthermore, the paper defines the idea of scalability, resource optimization, and integration with existing security systems.

  In the paper, Different NIDS solutions in cloud environments are explained such as Anomaly-based Detection, Signature-based Detection, and Hybrid detection. It also discusses the benefits of each detection technique and goes into detail about how the Hybrid Detection approach combines both Anomaly-based and Signature-based methods and has high levels of accuracy. Also, The paper discusses Regulatory regulations that are used to protect the data against attackers. These regulations range from Encryption to not sharing your password with anyone. Furthermore, the paper explains Future Directions and challenges for NIDS in cloud computing environments. One of the Future directions is using AI and machine learning techniques in NIDS to improve the ability to detect new attacks and threats. Furthermore, a comparison table was done between Anomaly-based, Signature-based,

and Hybrid detection with specific criteria like True Positive Rate and False Positive Rate. Finally, a data analysis was done to analyze the qualitative data obtained from the literature review.

The main result of this article is to discuss NIDS mechanisms explicitly created for security in a cloud environment. Also, due to the constantly changing cloud architecture, old-style security models must be more robust to tackle new security threats. For this reason, it has become essential to use anomaly-based, signature-based, and emerging behavior-based threat detection systems, as they help organizations boost their data security in the cloud. Considering information from recent studies, the article seeks to equip stakeholders with a concise overview of limitations. In my opinion, the paper is strong according to its claim. The paper aims to present a comprehensive Comparative Analysis of Approaches in Network Intrusion Detection in Cloud Environments.

2.2 Anomaly-based Intrusion Detection System

- This article [4] offers a comprehensive look into IDS research, catering to the needs of both researchers and practitioners. It covers a wide range of topics, from intrusion detection techniques to attack forms, relevant features, popular datasets, challenges, and potential solutions. The discussion dives deep into IDSs, exploring their requirements and performance metrics, and introduces a taxonomy based on information source, detection strategy, mode, and architecture. It provides a detailed analysis and comparison of network intrusion detection methods, focusing on anomaly detection techniques. Additionally, it presents a classification of computer network attacks, discussing their forms and relevant network traffic features for detection, as well as commonly used evaluation datasets. Finally, the article sheds light on research challenges and possible solutions.

  This paper claims that in contrast to recent surveys that only concentrate on anomaly-based intrusion detection techniques dataset-related issues or both. Our article fills this gap by delivering an updated survey of IDSs. Compared to existing surveys, our work stands out in several ways: it discusses IDS requirements, presents a thorough and structured survey of NIDSs, particularly those employing anomaly detection techniques, evaluates their strengths and weaknesses, introduces a taxonomy of IDSs based on four criteria—data source, detection strategy, mode, and architecture. For each criterion, we explore multiple approaches and analyze their performance, covering IDS data sources, dataset types, architecture, and detection modes. We illustrate a taxonomy of different network attack forms and detail the relevant features for their detection, unlike most surveys that focus only on four network attack types: DoS, Probe, R2U, and U2R. Additionally, we summarize commonly used datasets for IDS benchmarking, highlight their issues, and compare them based on properties. We present the performance criteria for evaluating IDSs and discuss recent research challenges along with potential solutions. Also,

none of the previously done surveys offer a comprehensive examination of intrusion detection approaches, types of IDSs, their requirements, available datasets, various attack forms, categories, tools, and relevant features for attack detection collectively.

The paper explains thoroughly IDS by explaining the Data source, Detection Strategy, Detection mode, and Architecture. IDSs can be categorized based on the source of information used to detect the intrusions: HIDS (Host-Based) and NIDS(Network-Based). The main difference between them is that HIDS monitors the traffic generated by the computer itself. This traffic describes the behavior of the users such as the accessed applications, opening files, and so on. On the contrary, NIDS examines the whole network traffic corresponding to the information exchanged between the computers in the network. in the Detection Strategy, we have two main types which are Anomaly-based and Signature-based. Also, in this paper, the two detection modes were explained which are offline and on-line IDS. There is also a comparison of intrusion detection strategies where we have criteria such as high detection rate, and low false alarm and we will see which IDS is better in these criteria. Added to that, the paper explains the different architecture strategies. There are two main strategies: centralized and distributed. A centralized IDS controls and analyzes the traffic data collected from the monitored network and reports the detection results using a central analysis component. In a distributed IDS, data monitoring and analysis are done at several locations, using several central analysis units. In Addition, the paper also discusses network attacks and provides us with a summary of network information-gathering tools such as Snoop. Also, a Summary of the datasets used for intrusion detection systems is evaluated.

This survey aimed to provide comprehensive insights into IDSs. Within this article, we present an updated survey focusing on NIDSs. We conduct a theoretical comparative analysis of various techniques, specifically providing an exhaustive examination of NIDSs employing anomaly detection methods, including their respective advantages and disadvantages. Furthermore, we delve into the requirements of IDSs, categorizing them based on four key criteria: data source, detection strategy, mode, and architecture. Distinguishing our work from others, we introduce a taxonomy detailing network attacks, associated tools, and the most crucial features for their detection. Our coverage extends to IDS data sources, dataset types, IDS architecture, and detection modes. Additionally, we offer a summary and comparison of available IDS datasets, along with performance criteria and identification of research challenges. Furthermore, the challenges outlined in Section 6 will be investigated in future works. In my opinion, this paper is strong as it delivered what it claimed. it provided us with a detailed review of IDSs and it presented a taxonomy of IDSs based on four criteria: information source, detection strategy, detection mode, and architecture. Also, it provided us with detailed information about attack forms.

- This article [8] offers a detailed survey of Anomaly-based Host Intrusion Detection Systems. It discusses the categories of intrusion detection systems which are

Host-based and Network-based. Also, it tackles the difference between Signature Based Detection and Anomaly Based Detection. Also, it states both the advantages and disadvantages of both detections. Added to that, the article goes deeply into anomaly detection discusses different Anomaly Detection Approaches, and explains each one in detail.

The aim of this paper is to offer a thorough and detailed overview of anomaly detection research. Within this overview, we will explore the strengths and weaknesses of various existing anomaly detection techniques. This survey focuses on the study of established anomaly detection methods, such as Machine Learning Detection, examining how these techniques can be adapted across different application domains. Also, it provides us with a proper understanding of various anomaly types and their merits.

The paper explains thoroughly what is IDS and the two main categories of intrusion detection systems (Host-based and Network-based). The paper mainly focuses on Anomaly based detection where it discusses four main types of Anomaly Based Intrusion Detection Systems. These types are Statistical Anomaly Detection, Data Mining Detection, Knowledge-Based Detection, and Machine Learning Based. Additionally, this paper focuses on discussing each approach and what are the advantages and disadvantages of using each one of them.

The primary aim of this paper is to present an overview of anomaly-based Host Intrusion Detection Systems (HIDS). Nowadays, HIDS holds increasing significance and is integral to most intrusion detection systems. This paper reviews various anomaly detection methods, assessing their strengths and weaknesses. It highlights that employing data mining algorithms and cluster-based approaches tends to yield more accurate results with fewer false alarms. Additionally, it suggests that hybrid solutions combining network-based and host-based HIDS can be more useful across diverse domains and that choosing which intrusion detection system to use is based on the organization's specific needs. This survey explores multiple approaches to formulating the anomaly detection problem. A detailed theoretical understanding of different anomaly types is essential for the development of robust intrusion detection systems. Finally, Future research should be directed into developing efficient anomaly detection systems that work with complex systems and facilitate real-time interaction among various components. In Conclusion, The paper in my opinion is strong as the paper offers a structured and comprehensive overview of anomaly detection techniques, highlighting their importance in developing robust intrusion detection systems. It covers a wide range of topics, including the review of existing techniques, comparative analysis, research practices, and identification of research directions.

- In [12] This paper gives us a comprehensive overview of anomaly-based Network Intrusion Detection Systems (NIDSs). The paper discusses cyber kill chain models and cyber-attacks that compromise network systems. Moreover, the paper describes various Decision Engine (DE) approaches, including new ensemble learning and

deep learning approaches. The paper also provides more details about benchmark datasets for training and validating DE approaches. Most of NADSs' applications, such as Data Centers, Internet of Things (IoT), as well as Fog and Cloud Computing, are also discussed. Finally, Challenges and future directions are mentioned at the end of the paper to aid future researchers.

This survey provides a detailed review of the Network Anomaly Detection System that gives a better understanding of designing anomaly detection in different domains.

The main contributions of this survey include the following:

- We provide a comprehensive discussion of network threats and intrusion detection properties.
- We describe an architecture for the Network Anomaly Detection System (NADS) with describing its components.
- We explain the recent methodologies, involving ensemble- learning and deep-learning algorithms, and challenges of designing an effective NADS.
- We conduct several experiments using different network datasets, feature selection, and DE techniques to demonstrate their applicability for evaluating NADSs.

The remainder of this paper discusses the components of IDS, evaluation metrices used for IDSs and challenges and future directions for NADSs.

The paper explains thoroughly what is Network Anomaly Detection System (NADS). It also discusses the network threats by having a comparison table of attack types and their examples. Additionally, the paper tackles the five main intrusion detection properties such as monitored environments, detection approaches, applications and deployments, anomaly types, and defense responses. The paper mainly focuses on Anomaly based detection where it discusses four main types of Anomaly Based Intrusion Detection Systems. Added to that, this paper focuses on comparing popular datasets such as KDD99 and NSL-KDD datasets and explaining both the advantages and disadvantages of using each dataset. Finally, we have important measures that could be used for estimating the efficiency and reliability of IDSs such as performance, Completeness, and Profile update.

The primary aim of this paper is to present an overview of Network Anomaly Intrusion Detection Systems (NADS). This study discussed the background and literature related to NADS. Due to rapid advances in technologies, computer network systems need a solid layer of defense against vulnerabilities and severe attacks. While an Intrusion Detection System (IDS) plays a crucial role in providing cyber security by adding a protective layer to ensure secure networking, it also encounters challenges in being developed to operate effectively in real-time and adaptable environments. This paper explores anomaly detection methodologies capable of

efficiently identifying both known and zero-day attacks. Transitioning from a Misuse Detection System (MDS) to a Network Anomaly Detection System (NADS) has been a challenge within the computer industry. However, this obstacle can be solved by designing its architecture with a focus on data sourcing, pre-processing techniques, and employing Differential Evolution (DE) mechanisms. Finally, a comparative table was made between popular datasets with their merits and demerits. In my opinion, this paper is strong as it gives us a detailed overview of NADS.

- In [14] This paper offers a thorough examination of anomaly detection systems and hybrid intrusion detection systems from recent years to the present. Additionally, it delves into current technological advancements in anomaly detection, while pinpointing unresolved issues and challenges within this domain.

  The paper claims that anomaly detection systems, as a subset of intrusion detection systems, have the potential to be extremely effective in detecting both known and unknown attacks by modeling normal system/network behavior. The authors argue that while commercially available intrusion detection systems are predominantly signature-based and require frequent updates, anomaly detection systems offer a promising alternative by focusing on deviations from normal behavior rather than known attack signatures. The paper also emphasizes the importance of addressing technological challenges such as high false alarm rates and scalability issues to enable the widespread adoption of anomaly detection systems. It also highlights the need for continued research and development in the field to overcome these challenges and improve the effectiveness of intrusion detection systems in detecting and preventing cyber threats.

  The paper involves conducting a comprehensive survey of recent literature in the domain of anomaly detection. The authors aim to assess the ongoing work in the field and consolidate existing results by reviewing and analyzing various anomaly detection techniques proposed in the last six years.

  Key aspects of the methodology include:

  1. Reviewing recent literature: The authors gather and analyze recent research papers, articles, and studies related to anomaly detection to understand the current state of the field.
  2. Assessing ongoing work: By examining the existing literature, the authors aim to evaluate the progress and advancements made in anomaly detection techniques in recent years.
  3. Consolidating existing results: The authors aim to synthesize and present a comprehensive overview of the different anomaly detection methodologies, approaches, and technologies that have been proposed in the literature.
  4. Identifying open problems and research challenges: In addition to reviewing existing techniques, the authors aim to identify the open research problems and challenges in anomaly detection to guide future research efforts in the field.

Overall, the methodology of the paper involves a systematic review and analysis of recent literature to provide insights into the current trends, challenges, and advancements in anomaly detection techniques.

The paper provides an overview of various anomaly detection methodologies, including classification-based anomaly detection, clustering-based anomaly detection, and association rules combined with classification for anomaly detection. Also, The authors highlight the technological trends in anomaly detection and discuss the importance of modeling normal system/network behavior to detect both known and unknown attacks effectively. Added to that, The paper addresses challenges faced by anomaly detection systems, such as high false alarm rates, scalability issues, and the need for effective interoperability among different intrusion detection technologies. Overall, the paper appears strong due to its comprehensive survey, identification of open research problems, discussion of technological trends, emphasis on effectiveness, and acknowledgment of challenges. These factors contribute to the credibility and significance of the research presented in the paper.

- in [2] A structured and comprehensive overview of various facets of network anomaly detection was provided so that a researcher can become quickly familiar with every aspect of network anomaly detection. Also, present attacks normally encountered by network intrusion detection systems were provided. Added to that, existing network anomaly detection methods and systems were categorized based on the underlying computational techniques used. Within this framework, we briefly describe and compare a large number of network anomaly detection methods and systems. In addition, we also discuss tools that can be used by network defenders and datasets that researchers in network anomaly detection can use. We also highlight research directions in network anomaly detection.

  The claim of the paper is to examine the state-of-the-art in modern anomaly-based network intrusion detection. The paper emphasizes the classification and evaluation of Network Intrusion Detection Systems (NIDSs) based on detection strategies and evaluation datasets. It also presents various detection methods, systems, and tools while discussing evaluation criteria for testing their performance. Additionally, the paper outlines research issues and challenges for future researchers and practitioners in developing new detection methods and systems for evolving network scenarios

  The methodology of the paper involves providing a structured and comprehensive overview of various facets of network anomaly detection. Also, The authors categorize existing network anomaly detection methods and systems based on the underlying computational techniques used. They also describe and compare a large number of network anomaly detection methods and systems within this framework. Additionally, the paper discusses tools that can be utilized by network defenders and datasets that researchers in network anomaly detection can leverage. The authors also highlight research directions in network anomaly detection.

  In this paper, we have examined the state-of-the-art in modern anomaly-based network intrusion detection. The discussion has emphasized two well-known criteria

to classify and evaluate NIDSs: detection strategy and evaluation datasets. We have also presented many detection methods, systems, and tools. In addition, we have discussed several evaluation criteria for testing the performance of a detection method or system. A brief description of the different existing datasets and their taxonomy is also provided. Finally, we outline several research issues and challenges for future researchers and practitioners who may attempt to develop new detection methods and systems for the latest network scenarios. In my opinion, this paper is strong as The paper explores a wide array of topics in network anomaly detection, covering everything from the basics to advanced detection methods, systems, and tools. It organizes and compares existing methods and systems, discusses how to evaluate their performance, and suggests future research directions. Overall, it seems like a valuable resource for researchers and professionals alike who are interested in this field.

2.3 Signature-Based Intrusion Detection System

- In [6] The paper presents a taxonomy of techniques used to minimize false alarms in signature-based IDS, categorizing them based on their approach and effectiveness. The paper also compares its survey coverage with other existing surveys, provides insights into the limitations and advantages of each technique, and discusses the potential for future research in this domain. Additionally, the paper includes a review of commercial SIEM tools that incorporate some of these false alarm minimization techniques to improve their performance in handling security events and IDS alarms.

  This paper seeks to thoroughly examine and assess the methods utilized to reduce false alarms in signature-based Network Intrusion Detection Systems (NIDS). It also aims to categorize and evaluate different strategies proposed in existing literature for tackling the issue of false alarms generated by IDS systems, providing insights into the pros and cons of each approach. Furthermore, it emphasizes the significance of minimizing false alarms in improving the performance of intrusion detection systems, particularly in expansive network environments where the sheer volume of alerts can pose challenges for system administrators.

  The authors conducted a thorough review of relevant research papers, surveys, and commercial tools related to false alarm minimization in signature-based IDS systems. This review helps in identifying and categorizing different techniques used to address false alarms. Based on the findings from the literature review, the authors develop a taxonomy of false alarm minimization techniques in signature-based NIDS. This taxonomy categorizes the techniques into different groups based on their approach and effectiveness in reducing false alarms.

  The results of the paper provide a comprehensive overview of the current state of research on false alarm minimization techniques in signature-based intrusion detection systems, offering valuable insights into the effectiveness, applicability,

and potential advancements in this critical area of cybersecurity. in my opinion, the paper is strong as it aims to provide a comprehensive review of existing false alarm minimization techniques in signature-based Network Intrusion Detection Systems (NIDS), and it successfully achieves this goal.

- In [5] The research paper focuses on evaluating the effectiveness of Snort, a signature-based intrusion detection system, in detecting zero-day attacks. The research methodology also involves setting up an experimental architecture with physical and virtual hosts, using various services and software to simulate real-world scenarios. Exploits from the Metasploit Framework were utilized to test Snort's detection capabilities.

  The main claim of the study is to challenge the common belief that signature-based network intrusion detection systems (SNIDS), such as Snort, are unable to detect zero-day attacks. The research aims to empirically validate this claim by testing Snort's detection capabilities against a set of zero-day attacks and known attacks. Also, The study provides evidence that Snort is indeed capable of detecting zero-day attacks, although with a lower detection rate compared to known attacks. This challenges the notion that SNIDS are ineffective against zero-day threats and emphasizes the importance of continuous evaluation and improvement in intrusion detection mechanisms.

  The methodology of the paper involves an experimental architecture that includes two physical hosts and seven virtual hosts. One physical host contains the attacked virtual hosts and the NIDS (Network Intrusion Detection System), while the second physical host contains the virtual attacker host. The attacked hosts are configured with various services, and the NIDS is equipped with Wireshark and Snort 2.6. The study uses a set of attacks sourced from the Metasploit Framework to test the effectiveness of the NIDS. Pilot tests were conducted to ensure proper setup and virtual machines were restored to saved states as needed. The paper also describes the process of exploiting server software using Metasploit and how SNIDS can detect attack codes. Added to that, it further details the experiment architecture, the types of attacks employed, and how zero-day attacks were measured. It also provides insights into the setup, configuration, and testing procedures carried out to evaluate the effectiveness of the NIDS in detecting both known and zero-day attacks.

  The results of the study on signature-based intrusion detection for zero-day attacks using Snort can be summarized as follows:

  1. Snort demonstrated the ability to detect zero-day attacks, with a mean detection rate of 17% for the tested zero-day exploits.
  2. The overall detection rate for known attacks was higher, with a mean detection rate of 54%.
  3. The study highlighted that while Snort can detect zero-day attacks, the detection rate for such attacks is lower compared to known attacks.

4. Analysis of the zero-day detection by Snort suggested a conservative estimate of 8.2%.
5. The study also investigated the false alarm rates and evasion potential of the corresponding signatures used by Snort.

These results challenge the common belief that SNIDS, like Snort, are unable to effectively detect zero-day attacks and provide valuable insights into the detection capabilities of signature-based intrusion detection systems. in my opinion, the paper is strong as the claim was proven wrong By testing 356 severe attacks on Snort, the study provides empirical evidence to challenge the common belief that SNIDS cannot detect zero-day attacks. The results show that Snort is capable of detecting zero-day exploits, albeit with a lower detection rate compared to known attacks.

- In [3] The paper explores how well signature-based intrusion detection systems (IDS) can identify network attacks. It conducts two experiments to evaluate the detection accuracy and precision of three specific signature-based IDS setups. The authors also delve into the key elements of a typical IDS setup, including decoder and pre-processing, detection engine, and alert/logging functions. They underscore the significance of fine-tuning methods to minimize false positives, achieved by disabling certain rules in default rule sets. Additionally, the research stresses the continual necessity of monitoring and updating signature databases to bolster detection efficacy. Finally, the study offers valuable insights into enhancing the performance of signature-based IDS in both identifying and mitigating web attacks.

  The claim of the paper is that signature-based IDS can effectively detect web attacks, but there is a need for continuous monitoring, tuning, and updating of signature databases to enhance detection capabilities and reduce false positives. The research aims to provide insights into improving the performance of signature-based IDS in detecting and responding to web attacks, ultimately contributing to the advancement of cybersecurity measures in network environments.

  The methodology of the paper involves two main experiments to evaluate the detection rate and precision of three signature-based IDS with predefined configurations.

  1. **Experiment 1**: Detection Rate Evaluation
     - This experiment focuses on determining the detection rate (recall) against various public and in-house datasets of URI attacks.
     - The goal is to assess how effectively the signature-based IDS can detect and identify different types of web attacks.
  2. **Experiment 2**: Precision and False Alarm Rate Evaluation
     - This experiment aims to evaluate the precision and rate of false alarms generated by the signature-based IDS.
     - The evaluation is conducted using a real-life large dataset to assess the accuracy of the IDS in detecting web attacks while minimizing false positives.

The methodology also involves discussing the components of a typical signature-based IDS, such as decoder and pre-processing, detection engine, and alert/logging mechanisms. Additionally, the study emphasizes the importance of tuning methods to optimize the performance of the IDS by deactivating unnecessary rules and continuously updating the signature databases to adapt to evolving threats in web attacks.

The result of the study provides valuable insights into the detection capabilities of signature-based IDS in the context of web attacks and highlights the importance of continuous monitoring, tuning, and updating of IDS configurations to improve detection accuracy and reduce false positives. In my opinion, the paper appears to be strong in terms of its methodology and contributions to the field of cybersecurity as the study conducts two experiments to evaluate the detection rate and precision of signature-based IDS in detecting web attacks, providing valuable insights into the effectiveness of these systems. Additionally, the paper proposes an efficient method for reducing false positives by tuning the IDS configurations, which can significantly improve the accuracy of detection while minimizing false alarms. Furthermore, the paper addresses the importance of continuous monitoring and updating of signature databases to adapt to evolving threats in web attacks, emphasizing the need for ongoing optimization of IDS configurations. By highlighting the challenges and opportunities in enhancing the performance of signature-based IDS, the study contributes to the advancement of cybersecurity measures in network environments. Overall, the paper's methodology, results, and contributions suggest that it is a strong contribution to the field of intrusion detection systems and web attack detection.

- In [15] The paper discusses a signature-based approach for intrusion detection proposed by Bon K. Sy. The approach involves deriving statistically significant event patterns from user/system access data to create unique signatures. These signatures are compared with a reference model to detect anomalies and potential security intrusions. Also, the effectiveness of the approach is evaluated using a publicly available dataset containing user masquerade. Added to that, the research aims to focus on masquerader intrusion but is extensible to other types of intrusions like network intrusion. The study also shows promising results in comparative evaluations, and future research will explore incorporating intrusion signatures and discovering other intrusion types. The approach involves applying statistical inference techniques, such as Chi-square goodness of fit, to determine the presence of masquerader intrusion. Additionally, the research mentions using Receiver Operating Characteristic (ROC) curve analysis to evaluate the approach's effectiveness. The signature-based approach is unique in characterizing normal behavior through statistically significant association patterns and conducting two-way mutual comparisons for distance measurements, distinguishing it from other statistical anomaly detection techniques

  The paper claims that the signature-based approach for intrusion detection, proposed by Bon K. Sy, utilizes statistically significant event patterns derived from

user/system access data to create unique signatures. These signatures are then compared with a reference model to detect anomalies and potential security intrusions

The methodology in the signature-based approach for intrusion detection includes the following steps:

1. Deriving statistically significant event patterns from user/system access data to create unique signatures.

2. Comparing real-time access signatures with a reference ground model to detect security intrusions.

3. Applying a model discovery utility to derive probability reference models for users, capturing access signatures with minimal biased information.

4. Reproducing data columns of Unix commands by shifting data and applying a function to discover statistically significant patterns in test data blocks.

5. Deriving observed and expected counts for statistically significant patterns and applying statistical inference based on Chi-square goodness of fit to detect masquerader intrusion.

6. Calculating correct detection rate, false positive rate, and false negative rate based on the results of test data blocks and using ROC curve analysis to evaluate performance.

The methodology emphasizes the importance of capturing normal behavior through access signatures and utilizing statistical inference techniques to detect anomalies and potential security intrusions effectively.

The results of the signature-based approach for intrusion detection indicate promising outcomes in detecting masquerader intrusion. By deriving statistically significant event patterns and comparing access signatures with a reference ground model, the approach shows effectiveness in identifying anomalies and potential security intrusions. The methodology, which includes applying statistical inference techniques like Chi-square goodness of fit and evaluating performance using ROC curve analysis, demonstrates the ability to detect masquerader intrusion with reasonable accuracy. The research also evaluates the approach based on comparative performance using a publicly available dataset containing user masquerade. By deriving access signatures and reference models, the study aims to reason the existence of security intrusions by comparing real-time access signatures with the reference model. The methodology's focus on capturing normal behavior through statistically significant association patterns and conducting two-way mutual comparisons for distance measurements contributes to the approach's effectiveness in detecting anomalies and potential security threats. Overall, the results suggest that the signature-based approach for intrusion detection, proposed by Bon K. Sy, shows promise in effectively identifying masquerader intrusion and lays the foundation for further research in detecting various types of security intrusions using similar methodologies. In my

opinion, this paper is strong as the approach is well-structured, utilizing statistically significant event patterns derived from user/system access data to create unique signatures for detecting masquerader intrusion. The methodology involves comparing real-time access signatures with a reference ground model, applying statistical inference techniques, and evaluating performance using ROC curve analysis.

- In [9] The paper discusses the challenges of intrusion detection in cloud computing and proposes a Hybrid Intrusion Detection algorithm for improved security in private cloud environments. To elaborate, the algorithm consists of three phases: Authentication, Misuse Analysis, and Anomaly Analysis. It aims to detect both Anomaly and Misuse Intrusions effectively. The paper also highlights the drawbacks of existing techniques and emphasizes the need for a comprehensive approach to intrusion detection. Additionally, it mentions the use of neural networks for intrusion detection in grid and cloud computing environments, noting the trade-offs in terms of detection accuracy and training time. The conclusion suggests that the proposed algorithm shows promise in terms of efficiency and scalability, with plans for further enhancements and generalization to different types of cloud environments.

  The claim of the paper is that the proposed Hybrid Intrusion Detection algorithm can enhance security and performance in private cloud environments by effectively detecting and mitigating intrusions.

  The methodology outlined in the PDF file involves the development of a Hybrid Intrusion Detection System for private cloud environments. The algorithm consists of three main phases:

  1. Authentication Phase: This phase involves validating and registering users to ensure only authentic users access the system.

  2. Misuse Analysis Phase: Users are validated based on login credentials and physical address to detect misuse intrusions.

  3. Anomaly Analysis Phase: Incoming user requests are validated based on bandwidth to detect anomaly intrusions. If any invalid request is detected, an alert for intrusion detection is sent to the administrator.

  The algorithm is designed to detect both Anomaly and Misuse Intrusions effectively, aiming to enhance security and performance in private cloud environments. The flowchart of the Hybrid Intrusion Detection Algorithm is provided in the document for a clear understanding of the methodology.

  The paper discusses the proposed Hybrid Intrusion Detection Algorithm and its potential benefits for enhancing security in private cloud environments. The document also emphasizes the efficiency and scalability of the algorithm, highlighting its ability to detect both Anomaly and Misuse Intrusions effectively. Additionally, the conclusion mentions plans for further improvements to the algorithm by including more parameters for intrusion detection and generalizing it for different types of cloud environments. In my opinion, the paper is weak as a comprehensive

evaluation of the algorithm's performance and comparison with existing methods is needed to further validate the effectiveness of the proposed approach.

Chapter 3

Methodology

3.1 Literature Review

During the last decade, several surveys of Intrusion Detection Systems, Anomaly-Based Intrusion Detection Systems, and Signature-Based Intrusion Detection Systems have been conducted. The surveys discussed multiple key points such as the most common datasets used to evaluate IDS, IDS classifications, and comparisons between different detection tools. All the literature discussed so far, does not focus on the comparison between Anomaly-Based Detection System and Signature-Based Detection System. However, in this article, we discuss this issue and we analyzed both systems with their tools.

3.2 System Selection

3.2.1 Signature-based Intrusion Detection System Tools

Tool utilized for the experiment

Snort

In this section, we will discuss three main types of intrusion detection systems : Snort, Sguil and Suricata

- **Snort :**

  Snort stands out as a popular free and open-source Intrusion Detection System (IDS). Written in C, it's renowned for its ability to analyze network traffic and detect various cyber threats based on user-defined rules. Currently maintained by Cisco Systems, Snort offers compatibility with both Linux and Windows operating systems. This powerful tool operates in three primary modes:

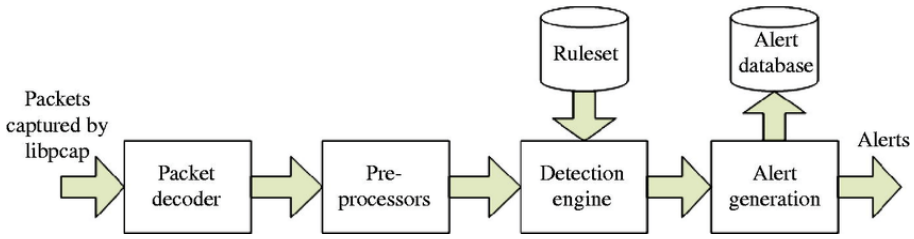

Figure 3.1: Snort Components
$https://www.researchgate.net/figure/Snort-architecture-Source-Snort-Manual_fig1_323003929$

- Packet Sniffing: collects and displays network traffic
- Packet logging: collects and logs network traffic into a file
- Network Intrusion Detection: Analyzes packets and compares network traffic against its rule set. if a match is found an alert will occur.

Snort rules are the core functionality of snort where these rules provide the instructions for Snort IDS to identify and respond to specific types of network traffic. if an alert occurs that means that there was a match between the network activity and predefined signatures. There are 3 types of rules set in Snort:

- Community rules: These are free rules that anyone can access. They offer protection against common threats and attacks.
- Snort Subscriber Rule Set (SRS): These rules are exclusive for subscribers. They offer broader coverage against threats and attacks, and they're regularly updated with the latest threats.
- Custom Rules: These rules are made by individual users to meet their specific needs.

Snort is logically divided into multiple components. These components work together to find any underlying threat or attack. These components are:

1. Packet Decoder
2. Pre-Processors
3. Detection Engine
4. Alert Generation

This figure shows how these components are arranged. Now we will explain briefly the task of each component.

Packet Decoder

The decoder is responsible for identifying the underlying protocols utilized in the packet (e.g., Ethernet, IP, TCP, etc.).

Pre-Processors

Pre-processors are essential elements or supplementary features in Snort, providing the ability to arrange or modify data packets before the detection engine assesses them for potential intrusions. Apart from preparing packets for analysis, specific pre-processors can also detect anomalies within packet headers, thus triggering alerts upon detecting this specific suspicious activity.

Detection Engine

Its responsibility is to detect if any intrusion activity exists in a packet. The detection engine employs Snort rules for this purpose. If a packet matches any rule, appropriate action is taken; otherwise, the packet is dropped.

Alert Generation

When the detection engine detects a packet that meets a predefined rule or shows unusual behavior, it produces an alert to inform administrators about a potential security issue. This alert usually contains details regarding the identified threat, including the attack type, source and destination IP addresses, timestamps, and any relevant information.

Advantages:

- Open-Source: Snort is a free and open-source detection system that eliminates licensing costs which can be expensive to people or organizations.
- Flexible: Snort offers a high degree of flexibility where you can define your own rules based on what you need or what threats are relevant to your organization.
- Real-Time Threat Detection:\textbf{ }Snort continuously monitors the network traffic and if any potential threat is found alerts will immediately occur. Therefore this will allow a quick response to security incidents, minimizing any potential damage.
- Powerful Traffic Analysis: Snort can monitor network traffic at a deep level, inspecting packets for malicious content, protocol anomalies, and suspicious activities. This deep inspection capability allows it to detect a wide range of threats, including malware.

Disadvantages:

- Rule Maintenance Burden: The effectiveness of Snort heavily relies on keeping its rule set up-to-date. New threats and attacks emerge constantly, so maintaining a comprehensive and accurate rule set requires ongoing effort.
- Potential for False Positives: Even with well-crafted rules, Snort can generate false positives, wasting analyst time investigating non-threatening events. Fine-tuning rules and understanding network traffic patterns are crucial for minimizing false positives.

- Complex Setup and Configuration: While Snort offers a powerful feature set, its initial setup and configuration can present a struggle, particularly for those new to network security.
- Its graphical user interface is not very user-friendly
- Incorrectly configured rules: The flexibility of creating custom rules in Snort comes with a trade-off. Inappropriate rule configuration can lead to:
  * **False positives:** Non-threatening events are mistakenly flagged as threats, generating unnecessary alerts and wasting IT resources.
  * **False negatives:** Actual threats going undetected because the rules haven't been properly defined to identify them.
- **Sguil :**

  Sguil stands for "Security Graphical User Interface for Linux". It is an open-source network security monitoring (NSM) tool which is a GUI . NSM is a tool used to analyze network traffic and generate an alert when an anomaly is found. Added to that, Sguil provides network security monitoring through the integration of different open-source tools such as Snort, Barnyard2, and other NSM.

  Advantages:

  * Enhanced Alert Analysis: Sguil goes beyond just displaying alerts from Intrusion Detection Systems (IDS) like Snort. It provides a user-friendly interface that allows analysts to Prioritize alerts. Doing that can help analysts sort through a flood of alerts by highlighting critical events and filtering out less important ones. This saves valuable time and ensures analysts focus on the most pressing threats.
  * Centralized Security Hub: Security teams often use a variety of tools for NSM. Sguil acts as a central hub, combining information from various security tools into a single interface.
  * Sguil acts as a central hub for network security monitoring, eliminating the need for multiple disparate tools. Without Sguil, administrators would require a collection of tools to achieve similar functionality, including: Vulnerability assessment tools to identify weaknesses in network defenses and Network traffic inspection tools to detect malicious activity.
  * Time: Sguil helps to reduce the intrusion detection time from five hours to 45 minutes, according to Vorant Network Security, Inc.

  Disadvantages:

  * Learning Curve: Sguil offers a wide range of features, which can be overwhelming for beginners. There's a learning curve involved in understanding how to effectively use all its capabilities for optimal threat detection and analysis.

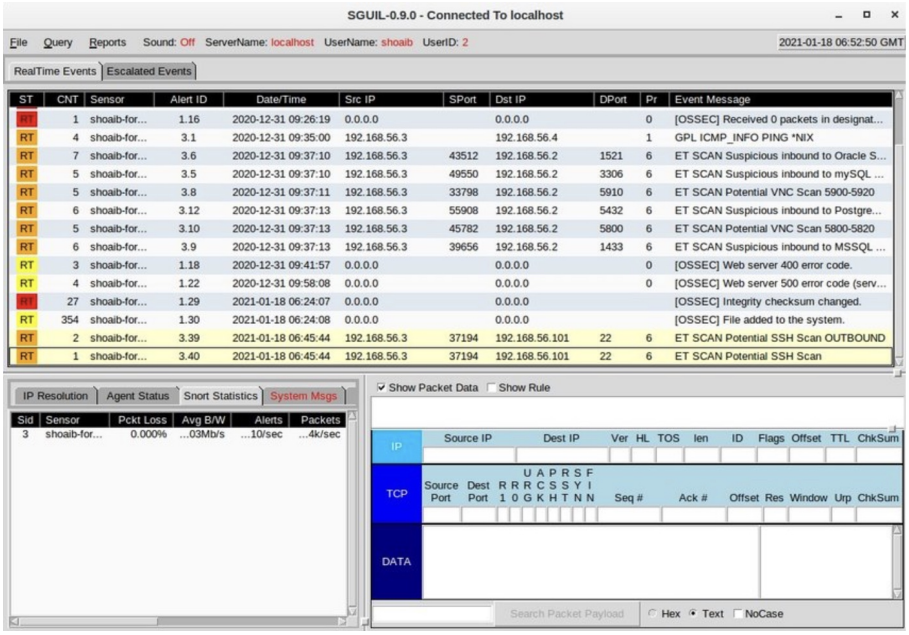

Figure 3.2: Example of Sguil interface
https://www.researchgate.net/figure/Sguil-interface-shows-the-alert-for-attacks-Squert-on-the-other-hand-uses-metadata-and_fig7_359684597

– **Suricata**

Suricata is an open-source network security monitoring engine. It functions as both an Intrusion Detection System (IDS) and an Intrusion Prevention System (IPS). The difference between IPS and IDS is that IDS only inspects the network traffic and then alerts specific security organizations if there is a suspicious activity while IPS examines the network traffic and if an alert occurs it can take action and drop the malicious packets. Suricata works by observing traffic and issuing alerts whenever that traffic matches the Suricata signatures of a known threat. The basic functions of Suricata include:

* **Network Traffic Analysis:** Suricata monitors network traffic by examining packets at the network layer. It inspects the contents of packets, including headers, payloads, and protocols, to identify potential security threats and anomalies.
* **Intrusion Detection:** Suricata compares network traffic against a set of rules and signatures to identify known attack patterns.
* **Intrusion Prevention:** Suricata can act as an intrusion prevention system by actively blocking and preventing detected attacks. It can drop or reject malicious packets and generate alerts for further investigation.
* **Protocol Analysis:** Suricata knows various network protocols and can analyze their behavior for signs of suspicious or malicious activity. It can detect protocol violations, malformed packets, and other anomalies that might indicate an attack or a security issue.

* **File Extraction and Analysis:** Suricata can extract and analyze files transmitted over the network. It can recognize known malware signatures, detect suspicious file types, and perform protocol-specific file analysis to uncover potential threats.
* **Log Generation and Reporting:** Suricata produces detailed logs and reports regarding detected events, alerts, and network activity. These logs can be utilized for forensic analysis and incident response(IR). To explain further, IR is a strategic approach in an effort to quickly identify an attack and have better cybersecurity.
  Advantages:
* Easy to use: Suricata is easy to install and configure, and it comes with a detailed user guide.
* High performance: Suricata is a high-performance engine that can be utilized to monitor both wired and wireless networks.
* Wide range of detections: Suricata can detect a wide range of threats and intrusions, including malware, network intrusions, and data breaches.
* Open source: Suricata is an open-source tool, which means that it is free to use by anyone and can be easily customized.
* Resetting connections: Suricata can terminate established connections that look suspicious.

Disadvantages:

* False Positives: Suricata uses predefined rules and signatures to identify threats. Overly strict or outdated rules can lead to false positives, where legitimate traffic gets flagged as suspicious. This can create alerts and waste valuable security personnel time investigating non-existent threats.
* Performance Overhead: While Suricata is known for its speed, it can still consume significant CPU and memory resources, especially when dealing with very high-bandwidth networks. This might lead to the need to upgrade the hardware.

3.2.2 Anomaly-based Intrusion Detection System Tools

In this section, we will discuss three main types of Anomaly-based Intrusion Detection Systems : Zeek, Dark Trace and Vectra AI

Tool utilized for the experiment

Zeek

• **Zeek**

Zeek is an open-source and free network intrusion detection system and a network traffic analyzer utilized by many organizations. It was invented in 1995 by Dr Vern

Paxson. Also, not too long ago Zeek was named as BRO. Zeek is not an active security device, like a firewall or intrusion prevention system (IDS). Rather, Zeek acts like a "sensor" that quietly monitors network traffic. Zeek transforms network traffic into transaction logs and file content. Let's breakdown Zeek's most crucial functionalities:

Network Traffic Analysis

Zeek sits on a sensor (hardware, software, virtual, or cloud-based) and passively analyzes network traffic flowing through it. It can monitor traffic at a packet level, inspecting the data packets that makeup network communication.

Deep Packet Inspection

Unlike some tools that only look at headers or basic information, Zeek can dig deeper as it can analyze protocols (like HTTP, DNS, and SSL) to understand the content of network traffic, allowing it to detect any anomalies and suspicious activities that might be hidden within encrypted data.

Event Generation and Logging

Based on its analysis, Zeek generates logs containing detailed information about the network activity. These logs include important details such as: source and destination IP addresses, protocols used, data payloads, and timestamps.

Customizable Rule Sets

Zeek offers you the ability to define your own rules in order to, detect specific threats relevant to your network environment. This level of customization ensures Zeek can identify threats that might not be covered by generic rules.

Advantages:

- Deep Packet Inspection: Unlike some tools that only analyze basic information about the traffic. Zeek inspects the actual packet thoroughly in order to get crucial details about the protocols used and data payloads. These details can aid in finding threats that might be hidden or malicious packets.
- Threat Analysis: Zeek doesn't just detect threats; it provides detailed logs for security analysts to analyze. This information will aid analysts in understanding the attacker's behavior and how they operate. Also, These detailed logs will help us to pinpoint where the problem lies in our security.
- Flexibility and Customization: Zeek's free open-source allows users to customize their own rules based on their own needs. Therefore, these customized rules will detect threats relevant to their unique network environment.

Disadvantages:

- Learning Curve: Zeek's scripting language and configuration can be complex for beginners. Users may need time and effort to be able to write their custom rules in Zeek's complex language.

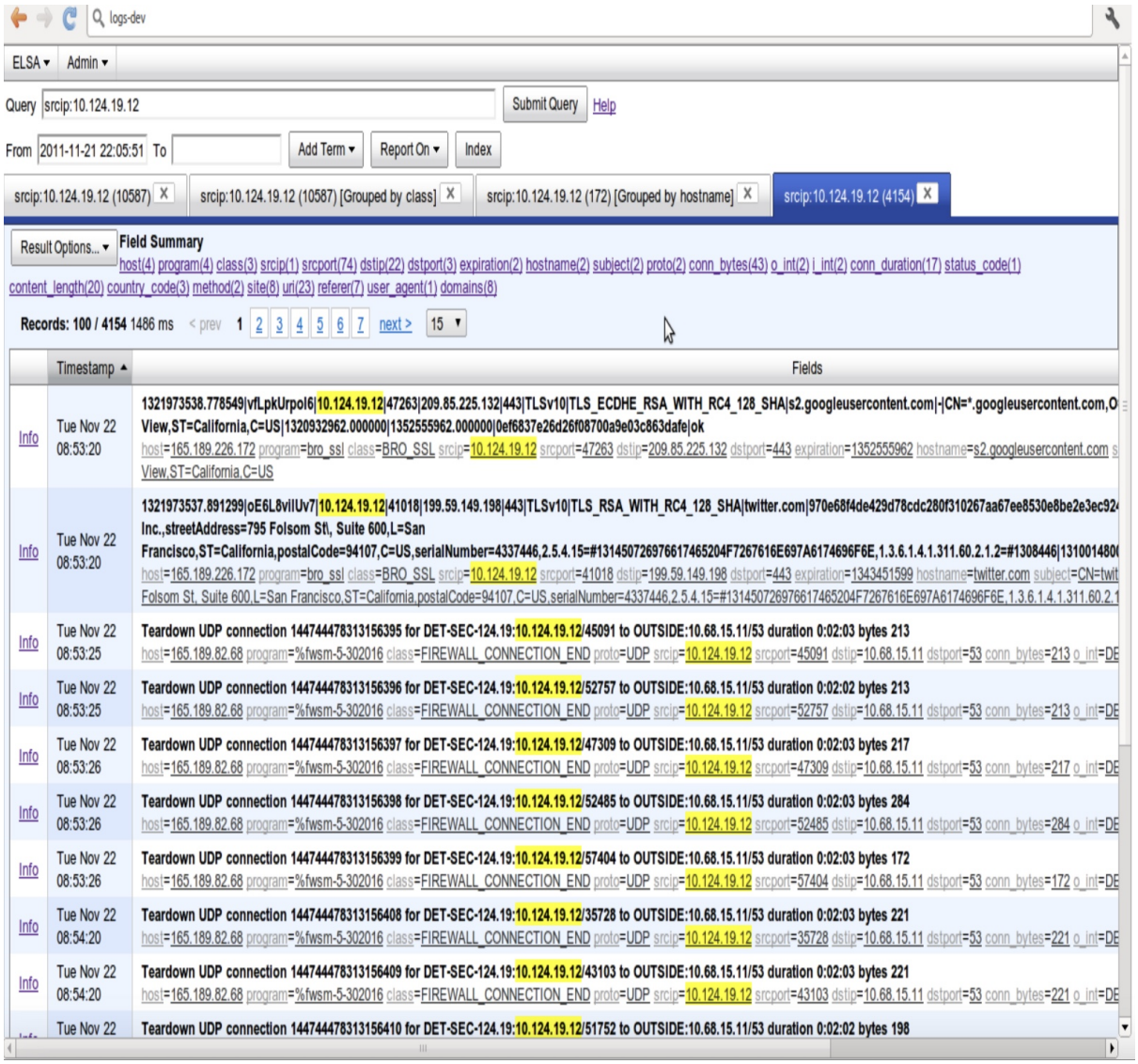

Figure 3.3: Example of Zeek Log
https://zeek.org/2012/01/04/monster-logs/

- Resource Consumption: Zeek can be resource-intensive, especially on older or less powerful hardware. Analyzing large volumes of network traffic can consume processing power and memory, potentially impacting overall system performance. This might necessitate deploying Zeek on dedicated hardware or using a distributed processing approach for very high-traffic environments.
- Lack of Real-Time Response: Zeek mainly focuses on network traffic analysis and passive monitoring. It doesn't have real-time response like other Intrusion Detection Systems.
- Maintenance Overhead: Like any network monitoring tool, Zeek needs continuous maintenance, updates, and rule refinement to adapt to new threats and attacks. This maintenance overhead can be demanding for smaller teams or organizations with limited money and resources.

• **DarkTrace :**

DarkTrace stands out as a global leader in artificial intelligence (AI) cybersecurity. Their technology utilizes self-learning AI to establish a deep understanding of your organization's typical network behavior within a week. This continuous learning process allows DarkTrace to effectively identify and respond to cyberattacks in real-time through a closed-loop AI system. This system operates in four distinct stages:. This AI loop consists of 4 Stages: *Prevent*, *Detect*, *Respond*, and *Heal*.

DarkTrace Prevent

This is the first stage in Darktrace's Cyber AI Loop. Its crucial aim is to stop threats before they even occur and this could be done by continuously monitoring your internal and external assets (servers, networks, applications) to identify potential weaknesses and misconfigurations that could be exploited by attackers.

DarkTrace Detect

This is the second stage in Darktrace's Cyber AI Loop. DarkTrace detect is the core component of the DarkTrace security platform as it continuously analyzes the network traffic and creates a baseline of the normal activity of your organization. With the knowledge of the normal behavior, DarkTrace can easily detect anomalies based on the deviations from the baseline. Upon detecting an anomaly, Darktrace provides security analysts with valuable context and insights. This includes visualizations, timelines, and other data points that aid analysts in determining if the anomaly is a genuine threat or a false positive.

DarkTrace Respond

This is the third stage in Darktrace's Cyber AI Loop. Darktrace's respond goes beyond just alerting security analysts about potential threats and anomalies. It can take pre-defined or on-the-fly automated actions to contain and disarm threats in real-time. Respond can immediately isolate devices or networks, to prevent the spread of any malicious activity, and reduce the amount of damage caused.

Figure 3.4: DarkTrace Cyber AI Loop
https://darktrace.com/blog/finding-the-right-cyber-security-ai-for-you

DarkTrace Heal

This is the final stage in the Darktrace's Cyber AI Loop. Darktrace Heal goes beyond static assessments as it can simulate real-world cyberattacks within your own organization's environment (in a safe, controlled manner) to test your security and spot any weaknesses. Based on these simulations, Darktrace Heal can then generate AI-powered incident response playbooks tailored to your specific network configuration and the attack scenario. On top of this, DarkTrace Heal provides real-time guidance to security analysts during any live cyberattack. After resolving the cyberattack, DarkTrace Heal generates a detailed comprehensive report regarding the date of the attack, the time of the attack, and the actions taken in that attack.

There are a variety of functions that DarkTrace offers such as:

1. **AI-powered Security Platform:** Darktrace's core offering is its self-learning AI security platform. This platform continuously monitors your network activity, user behavior, and other data sources. Using advanced machine learning algorithms, it analyzes this data to identify normal patterns of activity.
2. **Anomaly Detection:** Identifies anomalies that deviate from the established baseline, thus indicating threats and attacks.
3. **Focus on Unseen Threats:** Detects new or unseen attacks that normal security might miss.

Advantages:

- **Threat Analysis:** DarkTrace provides analysts with valuable insights
- **Automated Threat Response:** Since it relies on artificial intelligence, human intervention which could cause errors is minimized.

- **Proactive Threat Prevention:** With the help of DarkTrace Prevent, Attacks can be prevented before they happen.
- **Early Threat Detection:** Identifies threats before they cause serious damage.

Disadvantages:

- **Cost:** In comparison to traditional security maintenance and license, DarkTrace can be expensive for small organizations or startup companies.
- **Learning Curve:** There's a learning curve associated with getting the most out of this platform as DarkTrace needs skilled analysts to understand the data, configure it, and take appropriate action.
- **Not an open-source:** Unlike Zeek, DarkTrace is not an open-source therefore the source code can't be accessed or customized. This will lead to the inability of organizations to modify the source code according to their organizations' needs.

• **Vectra AI**

Vectra AI, a leading player in network detection and response (NDR), uses artificial intelligence (AI) to step up threat protection. Their AI-powered solution is powerful at quickly spotting questionable behavior and activities across your entire network, whether it's on-site or in cloud environments. Once something fishy is spotted, Vectra AI doesn't waste time; it quickly alerts your security team, so they can take action right away. Now, let's break down Vectra AI's key functions.

Anomaly Detection

By consistently examining network behavior in comparison to the established baseline, Vectra AI detects deviations that significantly diverge from the usual pattern. These deviations may suggest potential security risks or suspicious actions.

Machine Learning and Unsupervised Learning

Vectra AI uses machine learning algorithms such as unsupervised learning, to analyze the data from your network activity. This data can include network traffic, user behavior, cloud environment details, and more.

Detection of Unseen Threats

Vectra AI's AI capabilities excel at identifying unseen threats that traditional signature-based security solutions might miss. This is because it doesn't rely on pre-defined attack signatures, but rather focuses on anomalies in overall network behavior.

Advantages:

- **Early Threat Detection:** By prioritizing anomalies and analyzing overall network behavior, Vectra AI can detect threats before any potential damage can be done.

- **Reduced Response Time:** Due to faster recognition and investigation of threats this could lead to quicker response time thus minimizing any serious damage.
- **Adaptability to Evolving Threats:** The utilization of AI empowers Vectra AI to continually learn and adjust to emerging attack techniques, improving its capacity to identify unprecedented threats.

Disadvantages:

- **Cost:** In comparison to traditional security maintenance and license, Vectra AI can be expensive for small organizations or startup companies.
- **Integration:** For a more detailed view of your security landscape, Vectra AI might require integration with other security tools you already use such as Firewalls and other Intrusion Detection Systems.

3.3 Criteria Development

The criteria utilized in this survey to compare between Anomaly-Based Intrusion Detection System and Signature-Based Intrusion Detection System is Detection Accuracy , Speed and ease of implementation.

3.4 Empirical Analysis

3.4.1 Snort Tool

Before diving into the installation process of Snort, I initiated the deployment of Ubuntu. Ubuntu, distinguished by its robust Linux kernel foundation, stands as a perfect choice revered for its user-friendly interface, security protocols, and consistent updates. Selecting Ubuntu guaranteed a robust framework for Snort's functions. This initial setup not only ensured a stable and compatible environment for Snort but also provided a versatile platform for further customization and integration with existing network infrastructure.

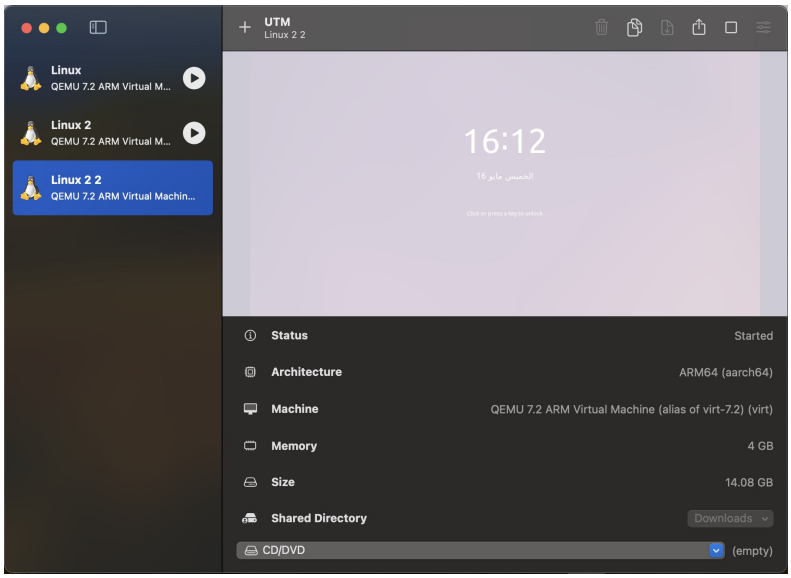

Figure 3.5: Ubuntu Interface

After that, I installed Snort on Ubuntu by using the following command line in the terminal:

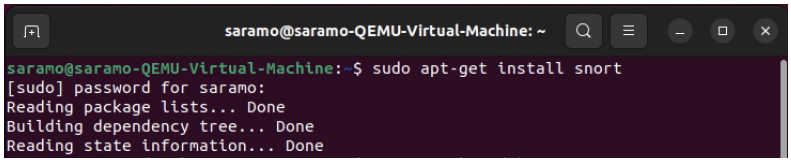

Figure 3.6: Snort Installation Command Line

Then, I checked my snort version after completing the installation using this command line:

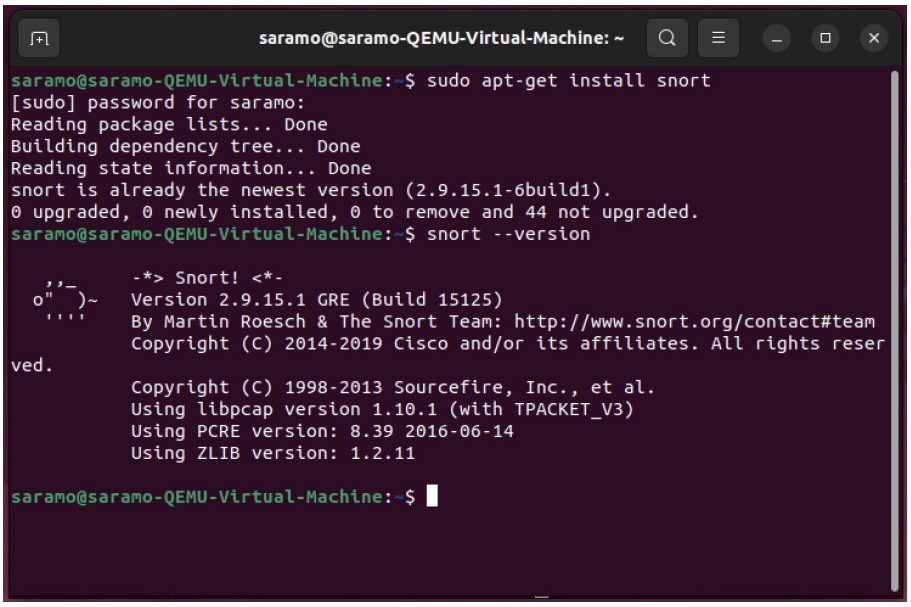

Figure 3.7: Snort Version

After that, I wanted to put snort rules which are utilized to enhance the effectiveness of the intrusion detection and prevention capabilities. Snort rules are essentially sets of instructions or patterns used by the Snort intrusion detection system (IDS) to detect specific types of network traffic or malicious activity. These rules are written in a specialized language that allows Snort to analyze network packets and compare them against predefined criteria.

The command line I used to open and edit the snort rules is:

Figure 3.8: Command Line used to edit or delete Snort Rules

After that local.rules file appeared on the screen. The local.rules file in Snort is a user-customizable file where you can define your own intrusion detection rules tailored specifically to your network environment and security needs. This file allows you to supplement the default rules provided by Snort with additional rules that are relevant to your organization's network. Added to that, the rules in local.rules file are typically given higher priority than the default rules provided by Snort. This means that they are evaluated before the default rules.

We can improve our network's intrusion detection capabilities by incorporating supplementary Snort rules from a trusted source. This GitHub repository (https://github.com/igbe/DNP3-Dataset-Plus-SnortRules), containing a curated set of Snort rules, offers a valuable resource to strengthen our existing defenses. Integrating these rules will broaden our detection scope against potential threats, ultimately enhancing our overall cybersecurity posture. Additionally, this repository serves as a helpful tool for maintaining and updating our customized rule set as security needs evolve.

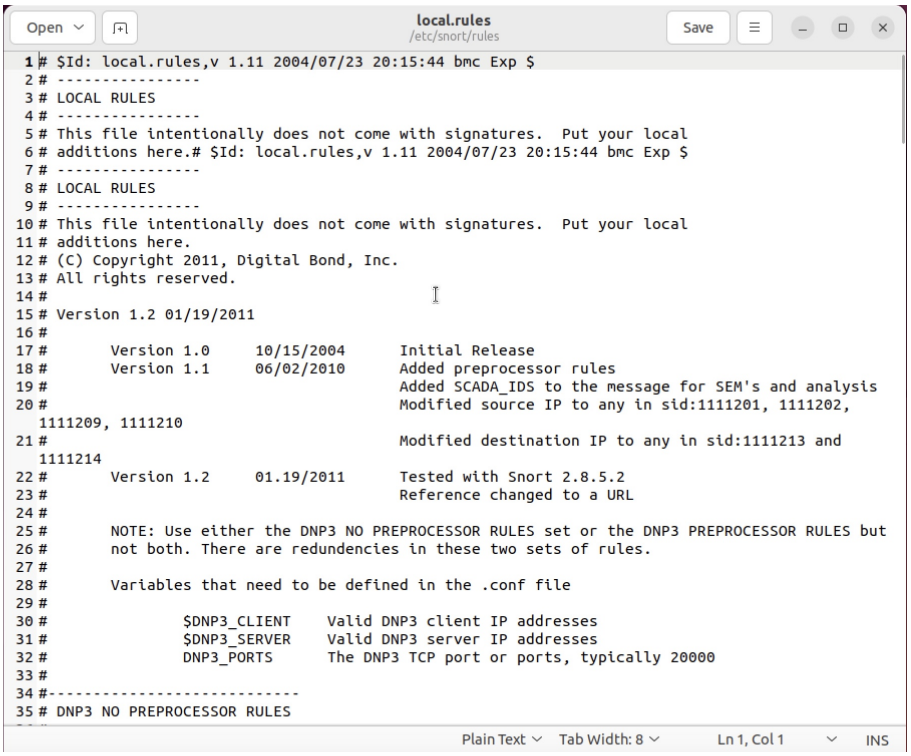

Figure 3.9: Local.rules File

Moreover, i analyzed the pcap file using this command line:

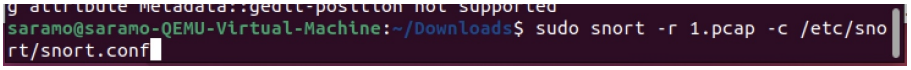

Figure 3.10: Analyze the pcap file

Afterwards, the details of the pcap file is found in the figure below

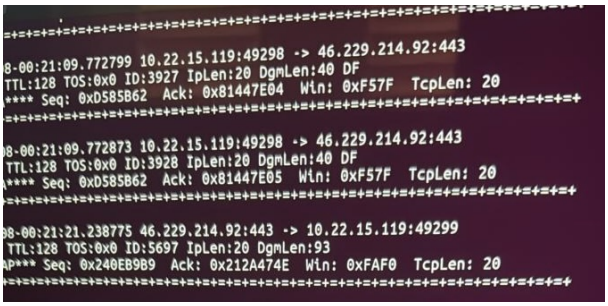

Figure 3.11: Details of the pcap file analysis

The results of the analysis is found below :

Table 1: Comparisons of popular datasets

PCAP File	Alert Count	Number of Hosts generating alerts
2018-11-13	9	4
2018-10-31	24	7
2018-08-12	3	2
2018-07-15	2	1
2018-05-11	1	1
2018-04-11	6	4
2018-03-10	1	1
2018-02-13	41	2
2018-01-16	31	11
2017-12-23	6	3

To find the average of Alert Count, you add them all together and then divide by the total count (which is 10 in this case):

(9+24+3+2+1+6+1+41+31+6)/10

24 / 10 = 12.4

3.4.2 Zeek Tool

On the other hand, I conducted multiple tests on various databases using an online pcap analysis. The specific tool employed can be found at https://dynamite.ai/. This test focuses on analyzing network traffic categorized as malicious or potentially constituting an attack. To elaborate, Dynamite.ai offers a comprehensive suite of tools for network traffic analysis, empowering security professionals to gain deep insights into network behavior and identify potential threats. Here's a breakdown of its functionalities and how they can benefit your organization:

- **Simplified PCAP Analysis:** Upload your PCAP files directly to the DynamiteLab Community platform. This free tier provides a user-friendly interface for analyzing network traffic captured within PCAP files.

- **Visualized Network Landscape:** DynamiteLab automatically generates a Network Graph, offering a clear visual representation of communication flows between devices on your network. This graphical depiction aids in understanding overall network activity patterns and identifying potential anomalies.
- **Timeline View for Activity Scrutiny:** The Timeline Analysis feature provides a chronological breakdown of network events within the PCAP file. This allows for a detailed examination of network activity at specific timestamps, facilitating investigation of suspicious events.
- **Communication Breakdown:** DynamiteLab delves deeper by providing detailed information about the communication protocols used, participating hosts, and other relevant elements within the captured traffic. This comprehensive analysis equips security professionals with a granular understanding of network communication patterns.

This experiment aims to evaluate network traffic using a dataset obtained from a trusted online platform: https://github.com/topics/malware-traffic-analysis. The specific dataset we will employ is the PCAP file located at https://github.com/topics/malware-traffic-analysis (2018-11-13-traffic-analysis-exercise.pcap.zip). We will begin by downloading this PCAP file for subsequent analysis with the chosen tool.

The Network Graph in the figure below provides a visual representation of the communication captured within the PCAP file. This graph offers valuable insights into the network activity and potential communication patterns. Here's a breakdown of the key elements:

- **Nodes (Blue Dots):** Each blue dot represents an IP address observed within the network traffic. These nodes symbolize the devices or hosts that communicated during the captured timeframe. The presence of 40 unique blue dots indicates that 40 distinct IP addresses participated in the network activity.
- **Links (Lines):** The lines connecting the blue dots represent the communication flow between the corresponding IP addresses. These links depict the source IP address (initiating the communication) and the destination IP address (receiving the communication). Analyzing the number and directionality of these links can reveal established communication patterns and potential network interactions. In this specific case, the presence of 39 links suggests a moderate level of communication complexity within the captured traffic.

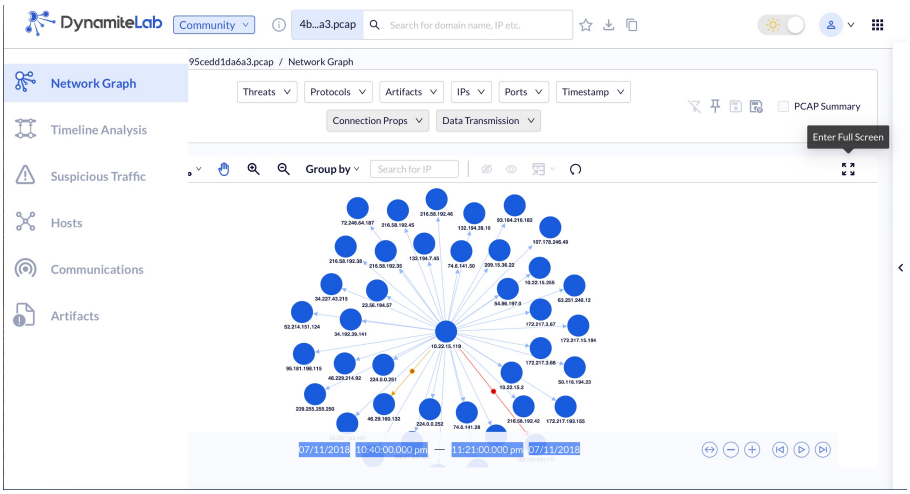

Figure 3.12: Network Graph

In the Network Graph, two distinct colored lines are observable: one in yellow and the other in red. These lines indicate alerts detected within these links. A yellow line signifies an alert of suspicious severity, while a red line denotes an alert of malicious severity

When clicking on the yellow link we will have these details :

1. Alerts info: This section indicates that an alert has been detected and is being reported to the system or network administrator for further investigation.

2. Severity: The severity level of the alert is described as "Suspicious." This suggests that the detected activity or event is not definitively malicious but warrants attention due to its potential implications or deviations from expected behavior.

3. Alert Count: The number of times this particular alert has been triggered is reported as "1." This indicates that the alert has occurred once within the specified timeframe or monitoring period.

4. Alert Categories: The alert is categorized under "Potential Corporate Privacy Violation." This suggests that the detected activity has the potential to compromise the privacy or confidentiality of sensitive corporate data, as discussed earlier.

5. Alert Signatures: The signature associated with the alert is identified as "ET POLICY PE EXE or DLL Windows file download HTTP." This signature indicates that the alert was triggered by network traffic consistent with a known pattern or behavior associated with potentially unauthorized or suspicious activity. Specifically, it suggests the download of Windows executable (EXE) or dynamic link library (DLL) files over the HTTP protocol, which may raise concerns about the introduction of potentially harmful software into the corporate network environment.

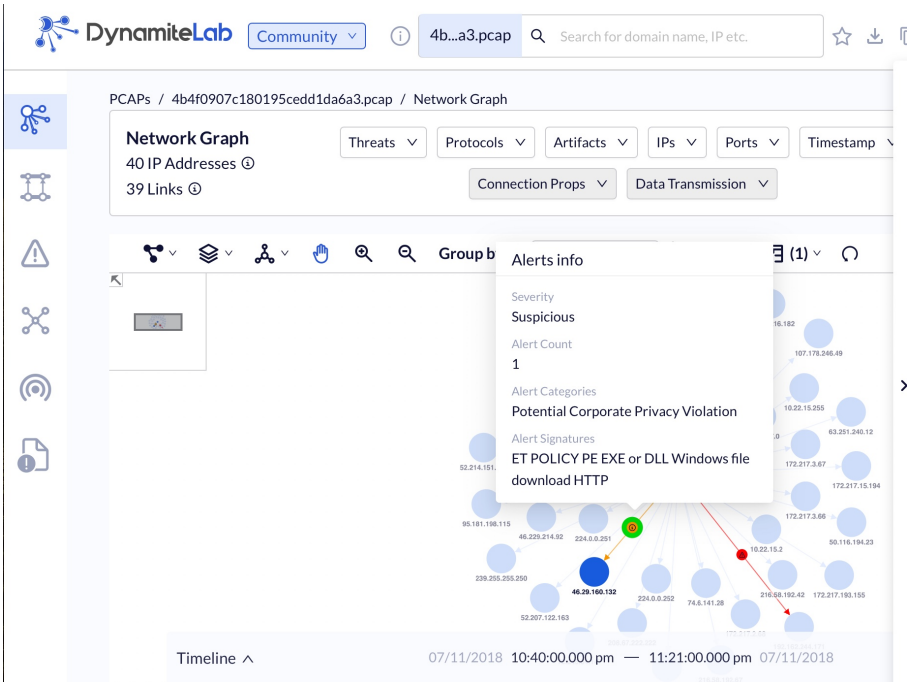

Figure 3.13: Yellow Link Alerts Info

Moreover, upon clicking the red link, the following details are revealed:

1. Alerts info: This section indicates that an alert has been detected and is being reported for further action by system or network administrators.

2. Severity: The severity level of the alert is categorized as "Malicious." This suggests that the detected activity is highly indicative of malicious intent, posing a significant threat to the security and integrity of the network.

3. Alert Count: The number of times this particular alert has been triggered is reported as "4." This indicates that the alert has occurred multiple times within the specified monitoring period, underscoring the importance of immediate attention and response.

4. Alert Categories: The alert is classified under "Malware Command and Control Activity Detected." This indicates that the detected activity pertains to communication or interaction with a command and control (C2) server associated with malware. Command and control activity is a hallmark of malware infections, as it enables threat actors to remotely control compromised systems and exfiltrate sensitive data.

5. Alert Signatures: The signatures associated with the alert are identified as "ET MALWARE Ursnif Variant CnC Beacon - URI Struct M1 (_2B)" and "ET MALWARE Ursnif Variant CnC Beacon - URI Struct M2 (_2F)."

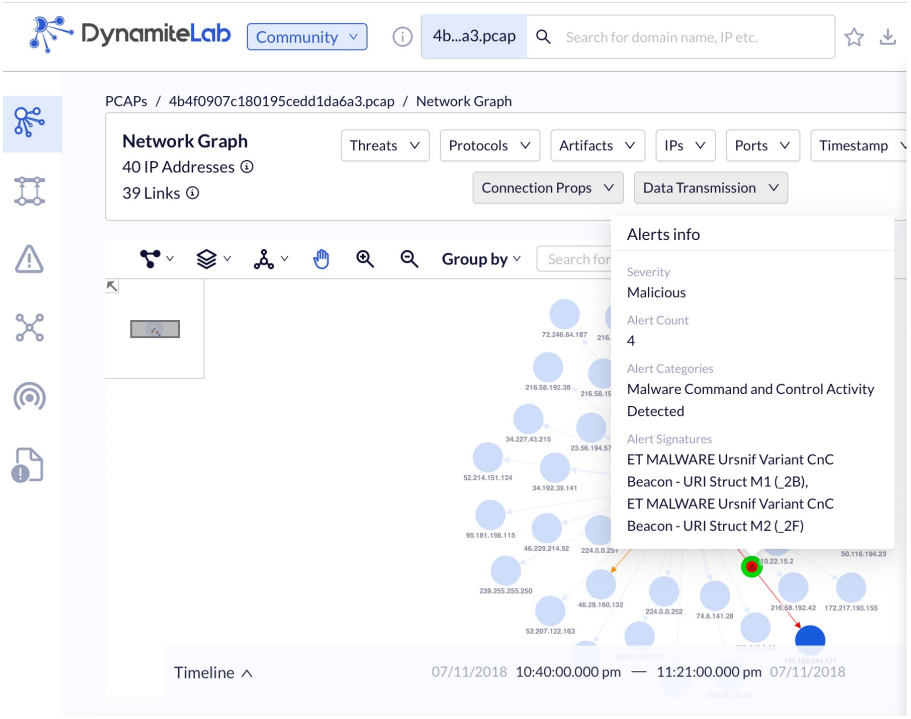

Figure 3.14: Red Link Alerts Info

We also have the Timeline Analysis Section which offers a chronological breakdown of network events captured within a PCAP file. This section provides a detailed view of network activity at specific timestamps, allowing for a more granular understanding of communication patterns and potential security incidents. For example, It has the time span of 39m 27s which is the amount of time between the first and last connections.

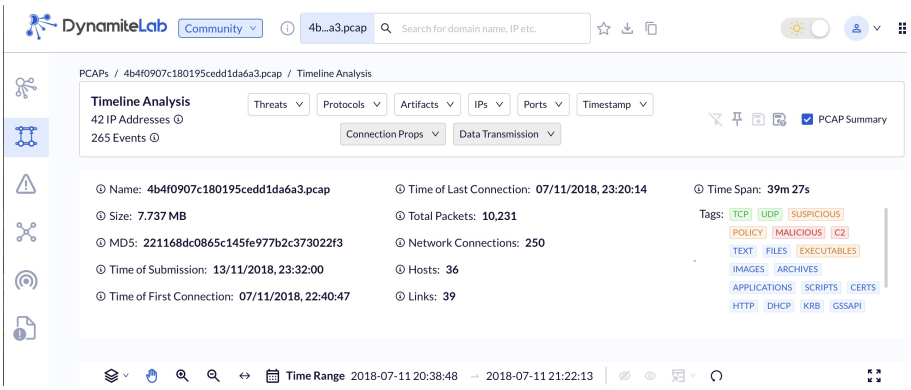

Figure 3.15: Timeline Analysis

Beyond the visualization tools, Dynamite.ai offers a critical section dedicated to identifying potential threats within the captured traffic. This section, often referred to as Suspicious Traffic provides valuable insights into suspicious activity:

- **Alert Count:** This section displays the total number of alerts generated during the analysis. Each alert signifies a potential security concern identified by Dynamite.ai.
- **Alert Source Identification:** The alerts typically pinpoint the IP addresses associated with the detected suspicious activity. This information helps security professionals prioritize investigation efforts by focusing on the sources responsible for triggering the alerts.

In this PCAP file we have 7 Alerts from 4 IP Addresses.

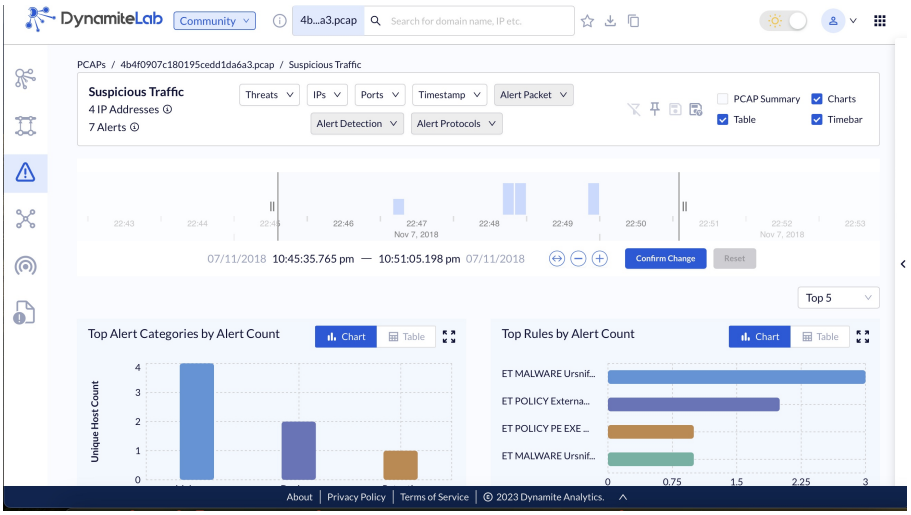

Figure 3.16: Suspicious Traffic

To summarize, I'll provide both a chart and a table displaying the distribution of alerts by IP Addresses.

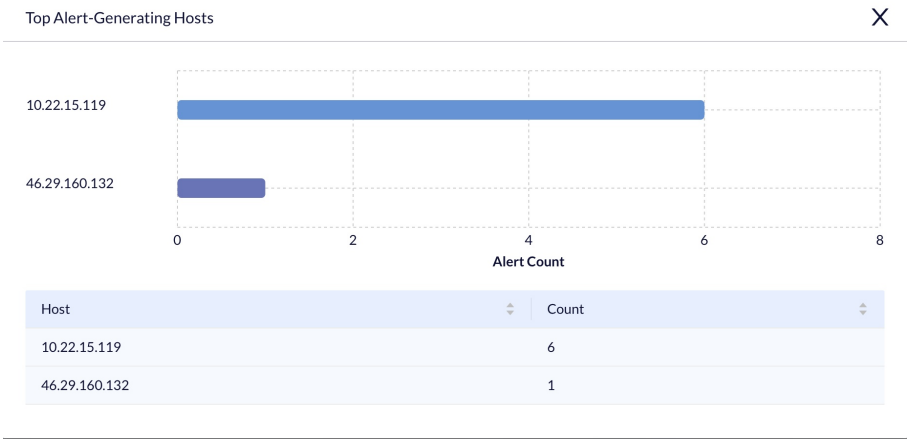

Figure 3.17: Comparison between Alert Count and Host

I'll conduct tests on various datasets and present the results, focusing on alert frequency, time duration, and the number of hosts generating these alerts. You'll find the outcomes detailed in the table below:

Table 2: Comparisons of popular datasets

PCAP File	Alert Count	Number of Hosts generating alerts
2018-11-13	7	4
2018-10-31	24	7
2018-08-12	1	2
2018-07-15	26	13
2018-05-11	24	3
2018-04-11	7	4
2018-03-10	28	9
2018-02-13	24	3
2018-01-16	0	0
2017-12-23	26	12

To find the average of Alert Count, you add them all together and then divide by the total count (which is 10 in this case):

(7+24+1+26+24+7+28+24+0+26)/10

=167/10

=16.7

So, the average of these numbers is 16.7.

To summarize the results, refer to the table below:

Table 3: Comparisons of Zeek and Snort

Tool used	Average Detection accuracy	Ease of Implementation	Speed
Snort	12.4	Hard	Faster
Zeek	16.7	Easy	Relatively slower than Snort

Chapter 4

Conclusion

In conclusion, neither Snort nor Zeek is inherently "better." The optimal choice depends on the specific requirements of the analysis. For this experiment, where deeper forensic analysis and potential discovery of unknown threats might be prioritized, Zeek's anomaly-based detection accuracy capabilities is a good fit. However, if real-time monitoring and efficient handling of large PCAP files are important, Snort could be a more suitable choice as it is faster. By carefully evaluating these factors alongside your specific requirements and technical expertise, you can select the intrusion detection system that best serves the purpose of your PCAP analysis and strengthens your overall network security posture.

Writing this thesis has presented several challenges. Firstly, explaining the complex technical concepts underlying both anomaly and signature-based intrusion detection systems has been difficult. These concepts relate to network security and data analysis, and making them clear and understandable, especially for readers unfamiliar with cybersecurity, has been a struggle. Moreover, I encountered significant challenges during the installation of Snort on my MacBook, primarily attributed to the complexity of its configuration language and command structures. This complexity necessitated a deep understanding of system architecture and networking protocols, adding layers of difficulty to the installation process.

For future work, it is imperative to conduct rigorous testing of both Anomaly-Based and Signature-Based Intrusion Detection tools across diverse datasets. This comprehensive evaluation will ensure the efficacy and reliability of our intrusion detection mechanisms. Also, We need to explore alternative evaluation metrics for comparing the performance of anomaly-based and signature-based IDS. Consider metrics beyond detection accuracy, Ease of implementation and speed such as false positive rate, false negative rate, and resource utilization, to provide a comprehensive assessment of IDS effectiveness. Added to that, we need to explore more the idea of combining both detection techniques to achieve the most robust security posture, considering a hybrid Intrusion Detection System (IDS) that merges anomaly-based and signature-based detection techniques. This approach leverages the strengths of both methods:

- **Signature-based detection** excels at identifying known threats quickly and accurately.

- **Anomaly-based detection** shines in uncovering novel and evolving attacks.

By combining these techniques, you can significantly improve overall detection accuracy while minimizing False Positives and False Negatives.

Furthermore, we need to explore the potential of advanced machine-learning algorithms for anomaly detection. Techniques like deep learning, reinforcement learning, or ensemble methods can significantly improve an IDS's ability to identify novel and sophisticated attacks with high precision. These algorithms can learn and adapt to ever-changing threat landscapes, providing a powerful layer of defense against emerging cyber threats.

Appendix

Appendix A

Lists

List of Figures

Bibliography

[1] Sina Ahmadi. Network intrusion detection in cloud environments: A comparative analysis of approaches. *Sina Ahmadi,"Network Intrusion Detection in Cloud Environments: A Comparative Analysis of Approaches" International Journal of Advanced Computer Science and Applications (IJACSA)*, 15(3), 2024.

[2] Monowar H Bhuyan, Dhruba Kumar Bhattacharyya, and Jugal K Kalita. Network anomaly detection: methods, systems and tools. *Ieee communications surveys & tutorials*, 16(1):303–336, 2013.

[3] Jesús Díaz-Verdejo, Javier Muñoz-Calle, Antonio Estepa Alonso, Rafael Estepa Alonso, and Germán Madinabeitia. On the detection capabilities of signature-based intrusion detection systems in the context of web attacks. *Applied Sciences*, 12(2):852, 2022.

[4] Suzan Hajj, Rayane El Sibai, Jacques Bou Abdo, Jacques Demerjian, Abdallah Makhoul, and Christophe Guyeux. Anomaly-based intrusion detection systems: The requirements, methods, measurements, and datasets. *Transactions on Emerging Telecommunications Technologies*, 32(4):e4240, 2021.

[5] Hannes Holm. Signature based intrusion detection for zero-day attacks:(not) a closed chapter? In *2014 47th Hawaii international conference on system sciences*, pages 4895–4904. IEEE, 2014.

[6] Neminath Hubballi and Vinoth Suryanarayanan. False alarm minimization techniques in signature-based intrusion detection systems: A survey. *Computer Communications*, 49:1–17, 2014.

[7] S Faizal Mukthar Hussain, R Karthikeyan, S Ramamoorthi, I Sheik Arafat, and S Syed Musthafa Gani. Enhanced protection for information and network using intrusion detection system. In *2023 2nd International Conference on Edge Computing and Applications (ICECAA)*, pages 280–286. IEEE, 2023.

[8] Shijoe Jose, D Malathi, Bharath Reddy, and Dorathi Jayaseeli. A survey on anomaly based host intrusion detection system. In *Journal of Physics: Conference Series*, volume 1000, page 012049. IOP Publishing, 2018.

[9] Roshan Kumar and Deepak Sharma. Signature-anomaly based intrusion detection algorithm. In *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pages 836–841. IEEE, 2018.

[10] Satish Kumar, Sunanda Gupta, and Sakshi Arora. Research trends in network-based intrusion detection systems: A review. *IEEE Access*, 9:157761–157779, 2021.

[11] Natalia Lewandowska. Intrusion detection systems: Categories, attack detection and response. 2024.

[12] Nour Moustafa, Jiankun Hu, and Jill Slay. A holistic review of network anomaly detection systems: A comprehensive survey. *Journal of Network and Computer Applications*, 128:33–55, 2019.

[13] Merve Ozkan-Okay, Refik Samet, Ömer Aslan, and Deepti Gupta. A comprehensive systematic literature review on intrusion detection systems. *IEEE Access*, 9:157727–157760, 2021.

[14] Animesh Patcha and Jung-Min Park. An overview of anomaly detection techniques: Existing solutions and latest technological trends. *Computer networks*, 51(12):3448–3470, 2007.

[15] Bon K Sy. Signature-based approach for intrusion detection. In *International workshop on machine learning and data mining in pattern recognition*, pages 526–536. Springer, 2005.