International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 02, February-2021

A Review and Classification of Intrusion Detection System in Data Engineering

Nilu Majeed
M.Phil. Research Scholar,
Department of Computer Science
Sree Narayana Guru College
Chavadi, Coimbatore, India

Dr. R. Priya
Associate Professor and HOD of Computer Science
Sree Narayana Guru College
Chavadi, Coimbatore, India

***Abstract*— Network revolution is an integral component of connectivity. The network risks have also increased from the latest developments of the internet. The conventional firewall methods are insufficient to survive the modern form of attacks from the internet. An effective intrusion is found to be associated with a machine weakness. Cyber threats are getting increasingly complex and this is making them difficult to track. In the future, that would be a big problem for intelligence forces, such as the security violation of data protection, integrity, and availability. Lots of Intrusion Detection System (IDS) have been proposed which can be generally categorized into Signature-based IDS (SIDS) and Anomaly-based IDS (AIDS) in the literature to protect against ransomware and cyber assaults. In this survey article, it provides a taxonomy of contemporary IDS information research, a thorough analysis of noteworthy recent works, and a summary of the data sets widely utilized for assessment purposes. The research provides a broad range of attacks and methods utilized by attackers to penetrate networks.**

***Keywords*— Security, Intrusion Detection System, SIDS, AIDS**

I. INTRODUCTION

The advent of malware and viruses presents a significant danger to the architecture of IDS. Malicious threats are getting increasingly complex and the key task is to reverse engineer malicious code, as malware writers employ obfuscation methods to avoid discovery by IDS

[1]. Non-deterrence has exacerbated security risks against internet users. Therefore, information protection has become a big issue in our everyday lifestyles [2].

There has been substantial influence from the zero-day attacks in multiple nations [3]. As shown in a 2017 study, three billion zero-day attacks were launched in 2016, and the frequency and severity of attacks were far greater than the previous years. In 2017, nine billion pieces of information were stolen from companies by hackers.

According to a Symantec study, security breaches are becoming popular. In the past, cyber attackers attacked mainly banks. It robbed bank accounts or confiscated credit and debit cards. Malware is now more ambitious because of which they are measuring their mettle against the banking sector themselves. Zero-day attacks are considered to be quite serious.

Cybercrime is growing as technological attacks are expected to get worse globally. Around the planet, there are a huge amount of cybercriminals motivated to hack information and unlawfully collect revenues.

Malware is purposely crafted to steal machine data and overtake networks. In 2017, the Australian Cyber Security Centre (ACSC) investigated the varying degrees of information used by the hackers. Hence the need to build a computer infrastructure to recognize and examine new, suspected malware. An intrusion detection system aims to recognize and prevent various kinds of attacks as quickly as possible, which conventional firewalls cannot accomplish. A continual need is a need for better IT protection [4].

There is a need for an up-to-date, comprehensive survey on the methods used for automated intrusion detection. There are a significant variety of similar works investigating the efficacy of spam filters. There is no straightforward response to the question of which strategies of data mining would be more successful. Secondly, the period required for constructing IDS is not included in the assessment of the efficacy of ‘on-line' IDS [5].

This article presents an up-to-date taxonomy, as well as a study of IDSs to date, as well as a classification of the proposed structures according to the taxonomy. The article provides viewers a description of critical parts of anomaly detection. This survey explores data mining methods as it pertains to intruder identification.

The remainder of the article is structured accordingly: Section 2 discusses a wide range of the latest IDS methods. Section 3 explores the methodologies of different IDS, the comparison of the study is given in Section 4 and the paper concludes in Section 5.

II. RELATED WORKS

In 2016 the authors in [6] identified an Intelligence Detection System which measures the overall detection efficiency of an IDS. An IDS must be able to communicate with various service layers within a device, such as a network interface, internet, transport, and application layers with a maximum achievable score of 4. Authors have analyzed seven intrusion detection programs against one standard metric to see how effective these approaches are in detecting intrusion attempts. Interaction capacity is a helpful predictor, however, current IoT intrusion detection attempts should be illuminated for review. Another essential aspect of IoT systems is that the

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 02, February-2021

absence of resource limitations does not impede the interaction capability of the device.

In 2017 the authors in [7] focused on the state-of-the-art of IoT protection research and outlined possible future research directions. Classification of current intrusion detection strategies was developed, which was focused on the detection process, IDS positioning technique, security hazard, and validation strategy. Since the authors have proposed a detailed method, however, this effort is not unique to the efficiency overhead of intrusion detection systems

In 2018 the authors in [8] examined the game theory and Markov decision processes for intrusion detection in IoT frameworks, demonstrating drawbacks of these methods. Authors analyzed how current intrusion prevention methods worked by comparing indicators such as True-Positives, False-Positives, True-Negatives, and False-Negatives.

In 2019 the authors in [9] proposed a model with the support of a specification-based methodology, to combat IoT protection problems. The authors claim that by utilizing a completely specification-based strategy, the proposed IDS achieves improved identification precision while ensuring greater security against previously unknown attacks. Although the authors have provided a thorough review of their methodology as well as the performance concerning identification precision, FPR, and FNR, however, they have not included metrics to show the performance efficiency of the proposed IDS.

In 2020 the authors in [10] have devised an intrusion detection system, which when coupled with the sinkhole attack will trigger further harm. Due to the specific relationship between various attack styles, we have researched the effect of multi-stage attacks on IoT infrastructures' security.

III. METHODOLOGIES

Intrusion is characterized as creating some amount of harm to the information system. Any intrusion that may impact the confidentiality/integrity/availability of information would be deemed a security breach. Activities that could wind up undermining the lawful usage of the computing resources are called an intrusion.

Forms of machine intrusions:

Cyber-attacks may be classified according to the form of actions and the nature of the goal. Every attack style can be grouped into one of the four following classifications:

Denial-of-Service (DoS) attacks are used to interrupt service or deter users from accessing the network.

The aim of a Probing-Attacks (PA) is to collect details about the network or computing device.

User-to-Root (U2R) attacks demand root access or administrator access to a single device or machine on which the attacker had low-level access.

Remote-to-local (R2L) attacks include transmitting packets to the victim computer, the "attacker machine." The cyber attacker knows about a person's actions and receives access to what rights the individual has on the machine.

An IDS is a hardware or software framework that senses suspicious behavior and warns security personnel of the intrusion. An IDS seeks to provide a behavioral study of harmful network traffic that cannot be detected by a conventional firewall. This is key to ensuring device availability and credibility, as well as the security of knowledge on computer networks. Intrusion detection systems can be roughly divided into two types: Signature-based IDS (SIDS) and Anomaly-based IDS (AIDS).

A. Signature-based IDS (SIDS)

The IDS based on the Signature is a type of intrusion detection that detects malicious software by matching patterns of known attacks. It was also termed as Misuse-Detection and Knowledge-based Detection. In SIDS, matching methods are used to determine a previous intrusion. As previously noted, when a signature matching one of the existing signatures is encountered, an alarm signal is triggered.
For SIDS, host activity logs are inspected to see what commands or actions have been previously exploited as malware. The primary objective is to create a database of intrusion signatures and to compare those signatures against current events and raise a warning when a match is triggered. For example, a rule in the form of “if: if (source IP address = destination IP address) then label as an attack” is a threat detection rule.

It provides a good feature detection for known intrusions. Even so, SIDS has difficulty in detecting zero-day attacks because no matching signature continues to exist in the dataset again until the signature of the new attack is retrieved and placed. Snort and NetSTAT are employed in numerous common tools, such as intrusion detection tools
Figure 1 shows the SIDS Architecture.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 02, February-2021

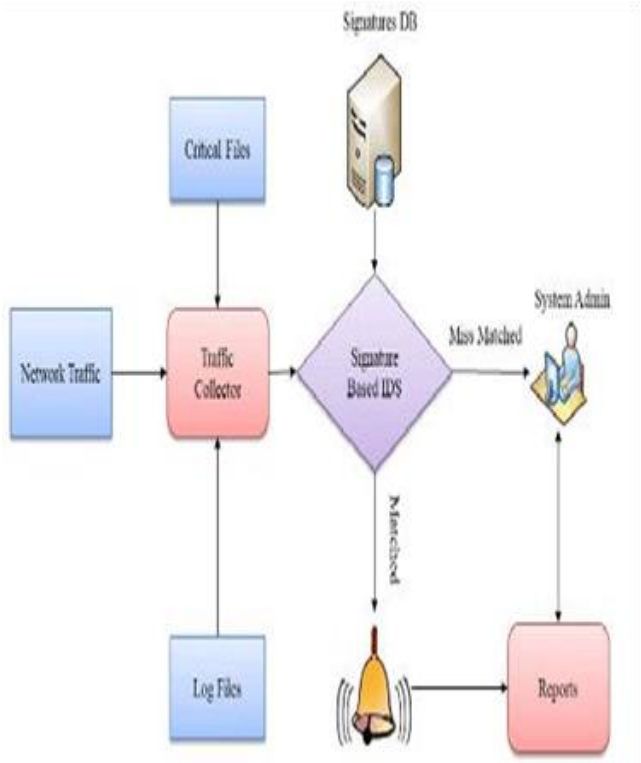

Figure 1: SIDS Architecture

Standard SIDS methods analyze packet data and aim to match a signature index. However, these methods cannot be used to identify large-spanning attacks. To make a signature of more sophisticated malware, signature information must be extracted over multiple packets.

This requires an information system to remember the content of earlier email messages. In particular there seem to be a variety of techniques in which Signatures are produced as state machines, structured string patterns, or semantic criteria for the creation of a signature for SIDS.

As the frequency of zero-day attacks has increased, the effectiveness of SIDS methods has decreased because they have no prior signature for zero-day attacks. The growth in the number of polymorphic variants of malware and targeted attacks undermines the sufficiency of this conventional approach. A good suggestion would be to utilize AIDS techniques. By profiling what is deemed acceptable behavior, instead of what is abnormal, a common standard can be established. The advantages and disadvantages of the SIDS have given in Table 1

Table 1: SIDS Advantages and Disadvantages.

Advantages	Disadvantages
• Very effective in identifying intrusions with minimum false Alarms (FA). • Promptly identifies the intrusions. • Superior for detecting the known attacks. • Simple design.	• Needs to be updated frequently with a new signature. • SIDS is designed to attacks for known signatures. When a previous intrusion has been altered slightly to a new variant, then the system would be unable to identify this new deviation of the similar attack. • Unable to detect the Zero-day attack. • Not suitable for detecting multi-step attacks. • Little understanding of the insight of the attacks.

B. Anomaly-based IDS (AIDS)

The effect of AIDS has attracted praise from many researchers since it overcomes the vulnerability of SIDS. A normal model of human activity has been developed utilizing mathematical, knowledge-based, or machine learning approaches. Every noticeable variance from the expected behavior is deemed a symptom of malfunction, and thus suggestive of a software malfunction.

The premise is that malicious acts are isolated from normal actions done by users. Abnormal activity peculiar to a certain user is known as an intrusion. AIDS research consists of two separate stages such as training phase and the testing phase. In the training phase of the experiment, a model is built of typical traffic behavior and then in the testing phase, a new data set is used to validate the model. AIDS may be categorized according to the type of teaching used, for example, statistical, knowledge-oriented, and machine learning.
The biggest benefit of AIDS is the potential to detect zero-day attacks so it doesn't involve signatures of irregular users' behaviors. It activates a threat alert for activities that are in disarray. AIDS includes various possible impacts. They have the potential to detect secret cyber-security risks. When an attacker begins making deposits through a compromised account which is not normal or flagged behavior for the customer, it activates an alert. Second, it would be impossible for a cyber-hacker to know what a common user activity is without finding secret details in the applications they use. Figure 2 shows the AIDS Architecture.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 02, February-2021

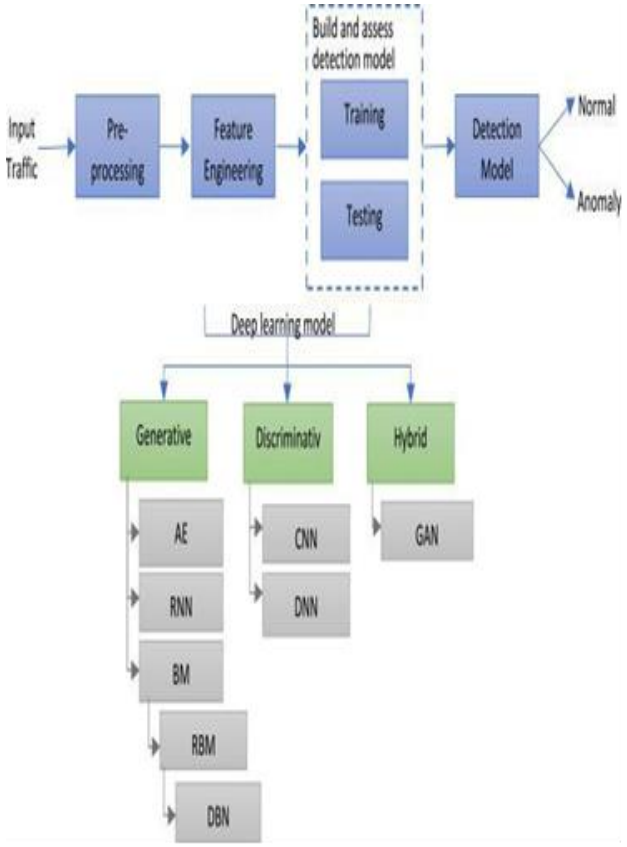

Figure 2: AIDS Architecture

SIDS identifies known attacks while AIDS detects "zero-day" attacks. Even so, the AIDS validation could even contribute to a large false-positive rate as these anomalies might be ordinary behaviors instead of genuine intrusions. The advantages and disadvantages of AIDS have given in Table 2.

Table 2: AIDS Advantages and Disadvantages

Advantages	Disadvantages
• Could be used to detect new attacks. • Could be used to create intrusion signature	• AIDS Cannot handle encrypted packets, so the attack can stay undetected and present a threat. • High false positive alarms. • Hard to build a normal profile for a very dynamic computer system. • Unclassified alerts. • Need initial training.

C. Host-based IDS (HIDS)

HIDS examines data that originates from the device it's built on and checks outlets such as the OS, WS, firewall, etc. It may identify threats that do not require network traffic. NIDS captures network data by collecting network packets via packet transfer, NetFlow, and other network data. Figure 3 shows the HIDS Architecture.

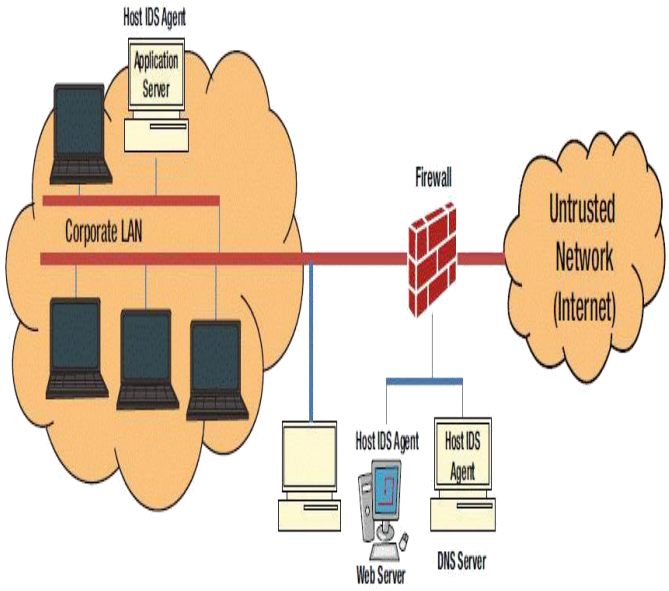

Figure 3: HIDS Architecture

The advantages and disadvantages of the HIDS have given in Table 3.

Advantages	Disadvantages
• HIDS can check end-to-end encrypted communications behavior. • No extra hardware required. • Detects intrusions by checking hosts file system, system calls or network events. • Every packet is reassembled. • Looks at the entire item, not streams only.	• Delays in reporting attacks. • Consumes host resources. • Needs to be installed on each host. • It can monitor attacks only on the machines where it is installed

D. Network-based IDS (NIDS).

NIDS can catch packets sent through several computers that are linked to a network. NIDS can track the external disruptive actions that may be launched by an external threat at an earlier point than the threats propagate. On the other side, NIDS has restricted abilities to inspect any of the data that travel across a high-speed transmission network because of the large amount of data being exchanged. Figure 4 shows the NIDS system architecture.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 02, February-2021

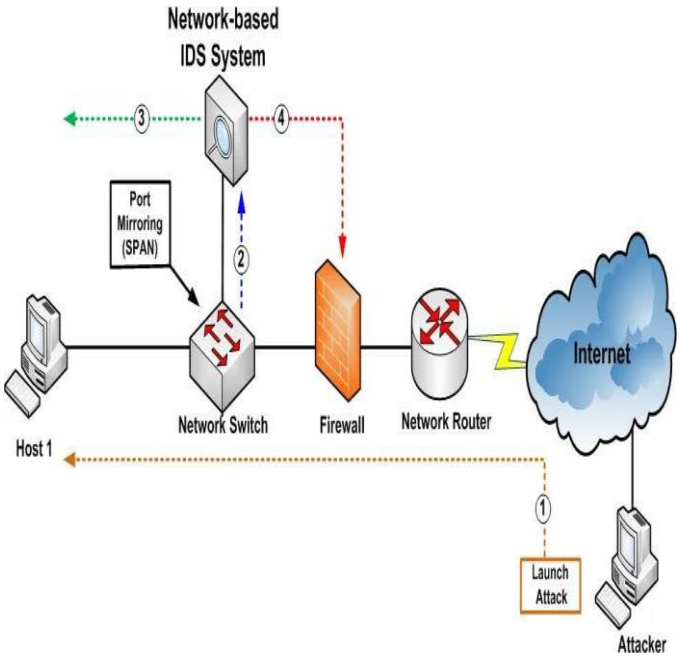

Figure 4: NIDS Architecture

The NIDS are installed at all critical locations within a network to provide a safe, robust, and multi-tier defense against both internal and external attacks. The advantages and disadvantages of the NIDS have given in Table 4.

Table 4: NIDS Advantages and Disadvantages.

Advantages	Disadvantages
• Detects attacks by checking network packets. • Not required to install on each host. • Can check various hosts at the same period. • Capable of detecting the broadest ranges of network protocols.	• Challenges is to identify attacks from encrypted traffic. • Dedicated hardware is required. • It supports only identification of network attacks. • The most serious threat is the insider attack.

IV. COMPARISION OF THE STUDY

Detecting threats concealed by prevention methods is a problem for both SIDS and AIDS. The identification and overcoming of avoidance methods will be calculated by the capacity of IDS to establish new and initial signatures to mask the modifications of the attacks. The effectiveness of IDS in detecting evasion strategies also needs further research. For instance, SIDS in standard representations could identify deviant patterns such as adding spaces, but they are still inadequate against a range of cryptographic algorithms.

Here it proves the AIDS is better for organizations that use encryption models while comparing it with SIDS.

A comprehensive IDS framework can support companies and have the ability to shield them against cyber-attacks. Even worse, existing intrusion detection strategies depend mainly on observing the applications behind computers. A critical detection strategy is needed to identify the zero-day and complex attacks until some information about the threat is obtained. Software and hardware intrusion prevention technologies may be combined to derive valuable functionality in all systems. The HIDS is only appropriate for limited scale activities where the company may only do a few events.

Here it proves the NIDS is better for a larger distributed organization that uses any type of Machine Learning model

V. CONCLUSION

Cyber attackers employ advanced tactics and approaches to obtain access to machines. And those that utilize cybercrime becoming more sophisticated and oriented. It is shown that computer criminals have proven their capabilities in concealing their identity, shielding their contact, and distancing themselves from illicit gains. Therefore, it is critical for computer systems to be secure against modern malware utilizing advanced intrusion detection systems. The above statement illustrates why programmers and app engineers ought to consider the benefits and weaknesses of existing IDS science. A review of intrusion detection framework methodologies, forms, and techniques was provided, along with its advantages and disadvantages. The paper discusses four different prevention methods to decide the ones that excel at hiding from the modern digital IDS. An efficient information protection management system must identify a variety of attacks accurately, even those which employ evasion techniques. To navigate around avoidance methods is a big problem for this field of research. According to this study, AIDS is suitable for the organization with the usage of encryption based models as their security and NIDS is suitable for the larger distribution organization with the advanced machine learning models.

REFERENCES

[1] Meng, G.; Liu, Y.; Zhang, J.; Pokluda, A.; Boutaba, R.Collaborative Security: A Survey and Taxonomy. ACM Comput. Surv. 2015, 48, 1:1–1:42. [CrossRef]

[2] Kumar, S.; Dutta, K. Intrusion Detection in Mobile Ad-hoc Networks: Techniques, Systems, and Future Challenges. Secur. Commun. Netw. 2016, 9, 2484–2556. [CrossRef]

[3] Nobakht, M.; Sivaraman, V.; Boreli, R. A Host-Based Intrusion Detection and Mitigation Framework for Smart Home IoT Using OpenFlow. In Proceedings of the 11th International Conference on Availability, Reliability and Security (ARES), Salzburg, Austria, 31 August–2 September 2016.

[4] Habibzadeh, H.; Nussbaum, B.H.; Anjomshoa, F.; Kantarci, B.; Soyata, T. A survey on cybersecurity, data privacy, and policy issues in cyber-physical system deployments in smart cities. Sustain. Cities Soc. 2019, 50, 101660. [CrossRef]

[5] Arshad, J.; Azad, M.A.; Mahmoud Abdellatif, M.; Ur Rehman, M.H.; Salah, K. COLIDE: A collaborative intrusion detection framework for Internet of Things. IET Netw. 2019, 8, 3– 14. [CrossRef]

[6] Gendreau, A.A.; Moorman, M. Survey of Intrusion Detection Systems towards an End to End Secure Internet of Things. In Proceedings of the IEEE 4th International Conference on Future Internet of Things and Cloud, Vienna, Austria, 22–24 August 2016.

[7] Zarpelao, B.B.; Miani, R.S.; Kawakani, C.T.; Alvarenga, S.C. A survey of intrusion detection in Internet of Things. J. Netw. Comput. Appl. 2017, 84, 25–37. [CrossRef]

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 02, February-2021

[8] Kiennert, C.; Ismail, Z.; Debar, H.; Leneutre, J. A Survey on Game-Theoretic Approaches for Intrusion Detection and Response Optimization. ACM Comput. Surv. 2018, 51, 90:1– 90:31. [CrossRef]

[9] Sharma, V.; You, I.; Yim, K.; Chen, I.; Cho, J. BRIoT: Behavior Rule Specification-BasedMisbehavior Detection for IoT-Embedded Cyber-Physical Systems. IEEE Access 2019, 7, 118556–118580. [CrossRef]

[10] Arshad, J.; Azad, M.A.; Abdeltaif, M.M.; Salah, K. An intrusion detection framework for energy constrained IoT devices. Mech. Syst. Signal Process. 2020, 136, 106436. [CrossRef]