Saarth E-Journal

Saarth E-Journal of Research

E-mail: sarthejournal@gmail.com
www.sarthejournal.com

ISSN NO: 2395-339X
Peer Reviewed
Vol.8, Issue.4 No.09

Impact Factor: 6.89
Quarterly
Oct. to Dec. - 2023

AN ANALYSIS OF INTRUSION DETECTION SYSTEMS: METHODOLOGIES, DATA SETS, AND OBSTACLES

***Dr. Vishalkumar Pradipkumar Patel**

ABSTRACT

As the level of sophistication of cyber attacks continues to rise, the responsibility of effectively identifying intrusions is becoming an increasingly challenging task. In the event that the security services are unable to prevent invasions, it is possible that the availability, integrity, and confidentiality of the data may be compromised. Regarding the protection of computers, there are a number of different intrusion detection systems available. "Signature-based intrusion detection systems (SIDS)" and "anomaly-based intrusion detection systems (AIDS)" are the two basic types of "intrusion detection systems". An investigation of regularly used datasets for evaluation purposes is included in this survey research. Additionally, a categorization system of modern "Intrusion Detection Systems (IDS)" is included, as well as a quick overview of relevant recent works. Furthermore, it elucidates the ways that attackers use to apply evasion strategies in order to remain undiscovered, and it examines the problems that lie ahead in finding and implementing effective solutions to these concerns, all with the goal of improving the security of computer systems.

KEY WORDS: Cyber Attack, Detection, Investigation, Strategies, Security.

(***Dr. Vishalkumar Pradipkumar Patel**, Asst. Prof. – Computer Science, Anand Mercantile College of Science, Management & Computer Technology. Email: vpp3188@gmail.com)

INTRODUCTION

An important obstacle that must be overcome in order for "intrusion detection systems (IDS)" to improve is the ever-changing nature of malicious software. As a result of malware developers employing a variety of techniques to conceal information and evade detection by an Intrusion Detection System (IDS), the identification of "unknown and obfuscated malware" has become a serious difficulty. Numerous security issues, including zero-day attacks that are directed at internet users, have been growing in frequency. The incorporation of information technology into our day-to-day activities has resulted in the heightened significance of computer security. The "United States of America and Australia" are two of the nations that have been heavily impacted by zero-day attacks. According to the "Symantec Internet Security Threat Report for 2017, it was discovered that there were over "three billion zero-day attacks that were reported in 2016". This represents a significant increase in both the number and severity of these assaults in comparison to the prior years (Symantec, 2017). According to the 2017 Data Breach Statistics (Breach_LeveL_Index, 2017), hackers have been responsible for the loss or theft of about nine billion data records since the year 2013. Instances of security breaches are increasing, according to the findings of a research conducted by Symantec. Symantec revealed in 2017 that in the past, fraudsters targeted those who had bank accounts or credit cards as their primary target audiences. According to Symantec (2017), new malware is growing more brazen by directly targeting banks in an effort to steal big quantities of money all at once. Because of this, the detection of zero-day attacks has emerged as a top priority in recent years. Instances of famous cybercrime that exemplify the quick worldwide spread of cyber threats reveal that even a little breach can impair a company's key services or facilities. This disruption can occur even if the breach is very minor. Cybercriminals are always looking for new targets, ways to illegally profit from their operations, and data to steal. They are constantly hunting for new targets. Malware is software that is constructed with the intention of causing damage to computer systems and exploiting vulnerabilities in "intrusion detection systems". An important research that was conducted in 2017 by the "Australian Cyber Security Centre (ACSC)" focused on the level of complexity of attackers. A "intrusion detection system (IDS)" that is both effective and efficient is required in order to identify new and complex forms of malware. With the exception of a conventional firewall, an "intrusion detection system (IDS)" has the potential to detect several forms of malware at an earlier stage. In tandem with the rising prevalence of

computer viruses, there has been a corresponding growth in the demand for enhanced "intrusion detection systems (IDSs)".

There is an immediate need for a comprehensive taxonomy and assessment of the present machine learning applications that have been used in intrusion detection throughout the course of the last several decades. Despite the fact that a number of studies have demonstrated that "Intrusion Detection Systems (IDSs)" have been developed by making use of the "KDD-Cup 99 or DARPA 1999 dataset," the efficiency of data mining techniques in comparison is yet unknown. Not only is the amount of time required to construct the intrusion detection system (IDS) an essential factor in determining the efficacy of online IDSs, but it is frequently disregarded when evaluating certain IDS strategies. A contemporary categorization is presented in this study, which also investigates the main research that has been conducted on "Intrusion Detection Systems (IDSs)" up to this point and makes use of this information to categorise the suggested systems. A detailed and well-organized analysis of modern "intrusion detection systems" is presented in this article. The purpose of this paper is to assist researchers in gaining a better understanding of the principles of anomaly detection. A description of the data-mining methods that are utilised in the process of "intrusion detection system design" is also included in this study investigation. We describe the ways that are based on signatures as well as those that are based on abnormalities (such "SIDS and AIDS"), as well as the many strategies that are utilised by each of these kinds of procedures. In this section, we will discuss a number of different approaches to AIDS and the methods that are utilised to evaluate them. After that, we will offer some recommendations for the most appropriate methods according on the type of invasion. The essay also examines the difficulties that are encountered by the "Intrusion Detection Systems (IDSs)" that are now in use. As opposed to earlier survey publications (Patel et al., 2013; Liao et al., 2013a), this study offers a comprehensive analysis of IDS dataset concerns that are significant to the research community in the field of "network intrusion detection systems (NIDS)". IDS stand for "intrusion detection system". "Intrusion Detection Systems (IDSs)" have not been thoroughly studied in previous research. This holds true in terms of the datasets, difficulties, and research methodologies that have been utilised. A comprehensive and up-to-date analysis of "intrusion detection systems" is presented in this article. Topics covered include datasets and methodology, among others. In addition, we discuss the difficulties that are associated with these systems and provide some potential solutions. IN previous years, many "intrusion detection studies" have been published. This research, along with those that came before it,

investigated the same “Intrusion Detection System (IDS) methodology and datasets”, as shown in Table 1. Both Axelsson’s taxonomy and survey on “intrusion detection systems (2000) categorise intrusion detection systems (IDSs)” according to the detection methods that they use. The research conducted by Debar and colleagues in the year 2000 focuses on the behaviour and expertise profiles that are associated with assaults. According to Liao et al. (2013a), “intrusion systems can be classified into five subclasses: statistics-based, pattern-based, rule-based, state-based, and heuristic-based”. An in-depth analysis was performed on each subclass. The aspects of “taxonomy, datasets, anomaly detection, and signature detection technique” are the primary areas of concentration for our research.

Survey	# of citation (as of 20/2/2024)	Intrusion Detection System Techniques						Dataset issue
		SIDS	AIDS				Hybrid IDS	
			Supervised learning	Unsupervised	Semi-supervised learning	Ensemble methods		
Lunt (1988)	219	✔	✖	✖	✖	✖	✖	✖
Axelsson (2000)	1039	✔	✔	✖	✖	✖	✖	✖
Liao, et al. (2013)	505	✔	✔	✔	✖	✖	✔	✖
Agrawal and Agrawal (2015)	108	✔	✔	✔	✔	✔	✔	✖
Buczak and Guven (2016)	338	✔	✔	✔	✖	✔	✔	✔
Ahmed, et al. (2016)	181	✖	✔	✔	✖	✖	✖	✔
This survey		✔	✔	✔	✔	✔	✔	✔

Review studies that have been conducted up till now have mostly concentrated on “intrusion detection technologies, datasets, different forms of computer attacks, and evasion approaches”. Despite the prevalence of “intrusion detection systems, dataset problems, evasion methods, and numerous forms of assaults”, there has been no systematic research that has adequately evaluated these topics. Due to the continuous growth of “intrusion-detection systems”, which justifies the need for diverse systems, it is also vital to have a system that is completely up to date. The purpose of this research is to present an updated taxonomy of the “intrusion-detection field” that improves upon the taxonomies that were previously developed by “Liao et al. (2013a)” and Ahmed et al. (2016) within the context of previous research. In respect to the discussion that was presented in the surveys that came before, this article focuses on the following:

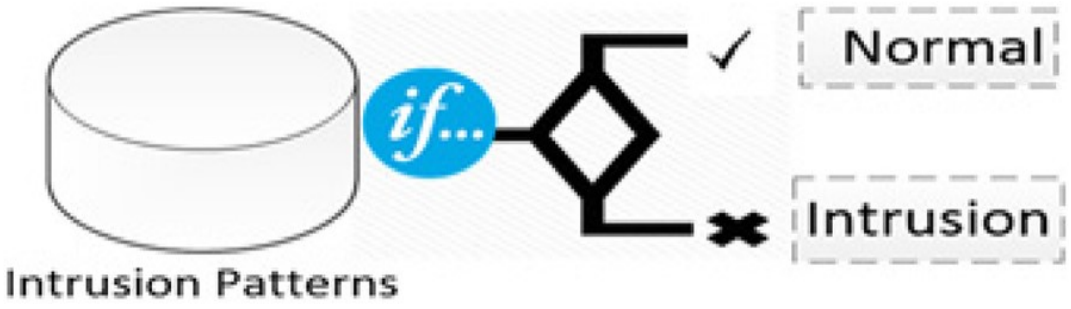
- Classifying various kinds of IDS with the key categories of assaults based on intrusion tactics.
- Discussing the significance of feature selection and presenting a categorization of network anomaly IDS assessment metrics.
- An analysis of the current IDS datasets addressing the difficulties of evasion methods.

INTRUSION DETECTION SYSTEMS

The term “intrusion” refers to any action that is carried out without authorization and that compromises an information system. Any attack that has the potential to compromise the availability, confidentiality, or integrity of information is referred to as an intrusion. It is considered an intrusion when actions are taken that render computer services unavailable to users who are authorized to use them. According to “Liao et al. (2013)”, the purpose of “intrusion detection systems (IDS)” is to maintain the security of computer systems by assisting in the identification and prevention of malicious activities. In situations where a conventional firewall is unable to recognize potentially harmful network activity or computer behaviour, an “intrusion detection system (IDS)” steps in to provide assistance during these situations. It is of the utmost importance to make certain that computer systems are adequately protected against any potential dangers that could compromise their “availability, integrity, or confidentiality”. There are two primary categories of “intrusion detection systems”, which are known as “Signature-based Intrusion Detection Systems (SIDS) and Anomaly-based Intrusion Detection Systems (AIDS)”.

SIGNATURE-BASED INTRUSION DETECTION SYSTEMS (SIDS)

Additionally referred to as “knowledge-based detection and abuse detection systems, signature intrusion detection systems (SIDS) are also known by their acronym”.



According to Khraisat et al. (2018), for the purpose of identifying known attacks, they use pattern matching. A prior intrusion can be identified by SIDS with the use of matching algorithms. An alarm will be generated in the event that an intrusion signature is found to be identical to a previous entry in the database. “Intrusion detection systems”, often known as IDS, examine host logs in search of potentially hazardous patterns of activities or instructions that have been discovered in the past. When it comes to “signature intrusion detection systems (SIDS)”, some studies refer to it as “Knowledge-Based Detection”, while others consider it to be “Misuse Detection” (Modi et al., 2013). An illustration of the theoretical operation of SIDS approaches may be seen in Figure 1. The creation of an accumulation of intrusion signatures, the comparison of the activities that are now being performed with that list, and the transmission of an alert in the event that a match is discovered are the key concepts. Take a look at the following directive: Utilizing a “if: antecedent -then: consequent” rule, classify an event as an attack if the “source of IP address” is identical to the “destination of IP address”. In their 2004 study, Kreibich and Crowcroft found that SIDS frequently provides a high level of detection accuracy for known breaks in security. It is difficult for SIDS to identify zero-day attacks before the “signature of the new attack” is separated and saved. This is because the database does not include a signature that corresponds to the particular assault. A number of commonly utilized tools that make advantage of SIDS include Snort, which was developed by Roesch in 1999, and NetSTAT, which was established by Vigna and Kemmerer in 1999. These are just two examples.

Typical approaches for “intrusion detection and prevention systems (IDS)” involve comparing incoming network packets with a database of signatures. By utilizing these methods, it is not possible to identify assaults that are persistent and include a “large number of packets”. When dealing with complex malware, it may be essential to ensure that signature data is obtained from a large number of packets. As a result, the “intrusion detection system (IDS)” has to remember information from previous packets. State machines, formal language string patterns, and semantic restrictions are some of the methodological approaches that have been proposed for the development of “SIDS signatures”. These are the ways that have been recommended the most often. SIDS approaches are becoming less effective as the frequency of zero-day attacks continues to rise (Symantec, 2017). This is because there is no previous signature for these attacks, which makes it difficult to detect them. When dealing with an increasing number of targeted assaults and polymorphic malware versions, this old technique may be much less effective than it already was. Putting into action the AIDS

methods that are described in the next part could be able to provide a solution to this problem. The operation of these methods is accomplished by studying and finding suitable conduct rather than concentrating on activities that are not typical.

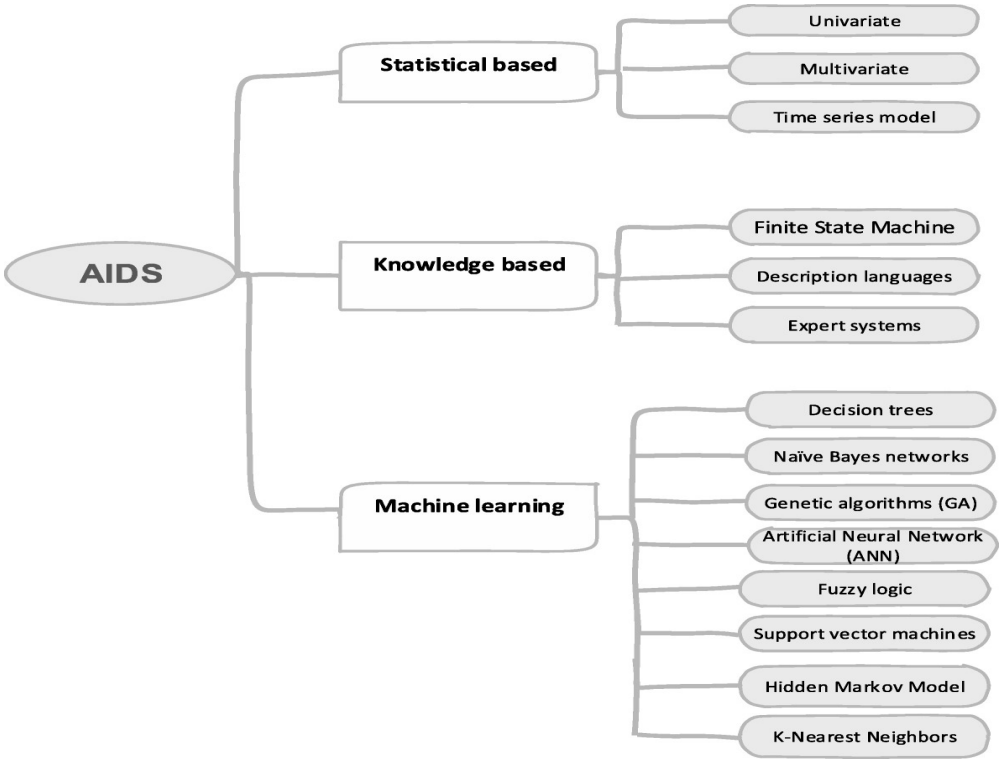
ANOMALY-BASED INTRUSION DETECTION SYSTEM (AIDS)

The potential of AIDS to provide a solution to the problem of sudden infant death syndrome is a source of fascination for researchers all over the world. The development of a typical computer system behavior model in AIDS is accomplished through the application of techniques that are based on machine learning, statistics, or expertise. An anomaly, also known as an incursion, is a behavior that significantly deviates from what is expected in relation to the model. The operation of these strategies is predicated on the assumption that dangerous actions are distinct from the actions done by regular users. Intrusions are behaviours that are exhibited by abnormal users and are marked by a deviation from the norm. Training and infection testing are the two stages that make up the development of AIDS. A model of typical behaviour is developed through the utilization of a typical traffic profile during the training process. Then, during "testing phase, a fresh data set is utilized" to demonstrate how well system can adjust to new and different types of intrusions. The classification of AIDS can be accomplished through a variety of approaches, depending on the training strategy that is utilized (Butun et al., 2014). These approaches include "knowledge-based, statistical, and machine learning-based knowledge-based techniques". According to Alazab et al. (2012), AIDS is able to recognize suspicious user activity without relying on a signature database, which enables it to detect zero-day attacks. The presence of AIDS is indicated by the activation of an alert that is triggered whenever abnormal behaviour is noted. AIDS offers a number of advantages, which serves as an additional incentive. The first thing they are able to do is identify any malicious activities that are taking place within the company. In the event that a hacker begins to conduct transactions in a compromised account that are not typical of the user's behaviour, an alarm will be triggered. In the second place, the one-of-a-kind profiles of the system make it extremely difficult for thieves to comprehend typical user behaviour without triggering an alarm. The use of Table 2 is one method that can be utilized to differentiate between "signature-based detection and anomaly-based detection procedures". AIDS is able to identify new and unknown attacks, whereas SIDS can only identify breaches that are already known to it. A significant number of false positives may be produced as a consequence of AIDS due to the fact that anomalies may simply indicate new behaviors that are typical rather than actual invasions.

Particular		Advantages	Disadvantages
Detection Methods	AIDS	➢ Might be able to spot more recent assaults. ➢ May find application in developing intrusion signature.	➢ Since AIDS is unable to decipher encrypted packets, the assault may remain unnoticed and pose a danger. ➢ A large number of false positives. ➢ A extremely dynamic computer system makes it difficult to construct a regular profile. ➢ Notified parties without classification. ➢ Requires first instruction.
	SIDS	➢ Promptly detects intrusions with minimal "false alarms (FA)". ➢ Finds the intruders quickly. ➢ Much better in spotting the recognized assaults. ➢ Design Simplicity.	➢ Regularly requires a fresh signature for updates. ➢ SIDS can identify attacks based on their recognised signatures. When an existing incursion has been slightly modified into a new variety, the system might not be able to detect this new variation of the same assault. ➢ Missing the zero-day assault. ➢ It is not designed to identify assaults that involve several steps. ➢ Familiarity with the assaults' insights is lacking.

METHODS FOR INTRODUCING AIDS

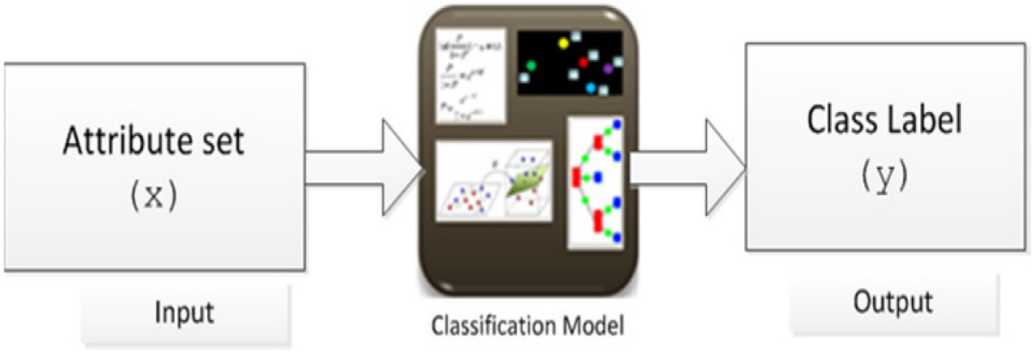
An overview of novel AIDS detection techniques that aim to improve detection accuracy and minimise the number of false alarms is provided in this section. AIDS techniques may be broken down into three primary categories: those that are based on statistics, those that are based on knowledge, and those that are based on machine learning. The process of developing a statistical model that is representative of common user activity involves gathering and evaluating all of the data entries included within a set of objects using a methodology that is driven by statistics. "In contrast to knowledge-based approaches, which try to infer desired actions from existing system data such as protocol specifications and network traffic patterns, machine-learning methods make use of training data to generate complicated pattern-matching capabilities". Figure 2 depicts the three primary categories, along with some instances of the subcategories that correspond to each of those broad groups.



INTRUSION DETECTION SYSTEM: SUPERVISED LEARNING

Here, many IDS supervised learning approaches are discussed in further detail. In addition to providing references to relevant academic material, this document provides in-depth descriptions of each technique. “Intrusion detection systems” that are based on supervised learning have the ability to identify intrusions by utilizing labelled training data. There are often two primary steps that make up a supervised learning process. These stages include teaching and evaluation. After the appropriate characteristics and classes have been discovered via the training phase, the system will acquire knowledge from data samples for further processing. The records that are included in supervised learning “Intrusion detection systems” are made up of pairs. One record identifies the data source, which might be a network or a host, while the other record provides the label, which is the output value, and indicates whether the data is normal or an intrusion. Feature selection is a method that may be utilized to get rid of features that are not essential. By applying certain features and a supervised learning approach with training data, a classifier is taught to recognize the link between input data and labeled output values. This is accomplished through the utilization of training data. There are many different supervised learning systems that have been investigated in the literature, and each of these systems has its own set of benefits and drawbacks. In the testing phase, the learnt model is utilized to classify unknown input into two distinct groups: the normal group and the intrusion group. Using a collection of feature

values, the classifier is transformed into a model that can make predictions about the type of data that is being received. The representation shown in Figure 3 illustrates a typical method for making use of categorization methods.

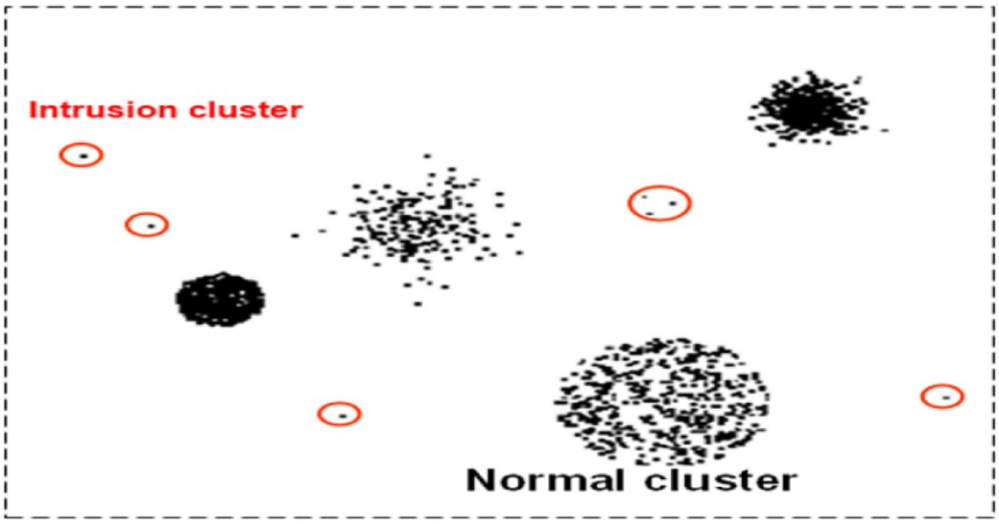


Some of the classification methods that may be utilized include |"decision trees, neural networks, naïve Bayes, support vector machines, rule-based systems, and nearest-neighbor". These are just some of the many methodologies that are available. For the purpose of constructing a classification model, each approach makes use of a different learning process. In order to be considered a successful classification algorithm, it must be able to handle the training data properly and accurately categorize records that are unknown. The fundamental objective of the learning algorithm is to facilitate the development of classification models that demonstrate a dependable capacity for generalization.

INTRUSION DETECTION SYSTEM: UNSUPERVISED LEARNING

An approach for machine learning known as unsupervised learning has as its goal the extraction of important insights from input datasets that do not include any class labels. The data points that are entered are frequently seen as a collection of factors that are entirely random. In the following step, the dataset is transformed into a joint density model. On the other hand, supervised learning makes use of output labels in order to teach the system to produce the necessary outputs for new data points. This is in contrast to unsupervised learning, which automatically organizes data into classes as it learns. Training the model with data that has not been labeled is an example of unsupervised learning, which is used in the context of the creation of "Intrusion detection systems". When compared to uncommon occurrences, the results of typical events need to manifest as substantial clusters. Therefore, any occurrences that form little clusters after the data has been sorted are deemed to be intrusions (Fig. 4 to be more specific). In addition, because of the characteristics that

distinguish them from one another, there is no correlation between aggressive incursions and benign happenings.



A substantial amount of research has been carried out on "cyber-physical control systems (CPCS)" that make use of unsupervised learning for the purpose of identifying threats and implementing reactive mitigation strategies. As an illustration, Alcara (Alcaraz, 2018) was the one who suggested a resilience strategy that was built on redundancy. Additionally, he suggested the implementation of a separate network sublayer for the purpose of handling context. This sublayer would make use of "data mining techniques such as k-means and k-nearest" neighbour in order to discern between different points of view. It would be the responsibility of this sublayer to routinely gather consensus data from the driver nodes that are under the supervision of the main network. The utilization of the Hybrid-Augmented device fingerprinting that was developed by Chao Shen and colleagues can be beneficial to the networks that are used for industrial control systems. Network packets were analyzed using a variety of machine learning techniques in order to identify aberrant activity and localize "intrusions in industrial control systems (ICS) networks".

ASSESSMENT CRITERIA FOR INTRUSION DETECTION SYSTEMS

Classification metrics for "Intrusion detection systems" abound, with certain measures going by more than one name. The following are some of the most common metrics used to assess IDS:

"True Positive Rate (TPR)": The ratio is calculated by dividing the total number of assaults by the number of attacks that were correctly predicted. An intrusion detection system having

a True Positive Rate of 1 while detecting all intrusions is uncommon. TPR can also be referred to as Detection Rate (DR) and Sensitivity. The true positive rate (TPR) can be represented by a mathematical equation

$$\text{TPR} = \frac{TP}{TP+FN}$$

"False Positive Rate (FPR)": The ratio of misclassified lawful cases as attacks to the total number of genuine instances is utilised for its determination.

$$\text{FPR} = \frac{FP}{FP+TN}$$

"False Negative Rate (FNR)": When a detector doesn't pick up on an abnormality and instead labels it as normal, it's called a false negative. A mathematical expression for the FNR is:

$$\text{FNR} = \frac{FN}{FN+TP}$$

"Classification rate (CR) or Accuracy": The ability of the IDS to differentiate between common and uncommon traffic patterns is measured by the CR. It is defined as the accuracy rate, which represents the fraction of correct predictions among all cases:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

CHALLENGES OF IDS FOR ICSS

There are two basic components that make up an Industrial management System "(ICS): the Supervisory Control and Data Acquisition (SCADA) hardware", which is accountable for the collection of sensor data and the management of mechanical equipment, and the software interface, which enables human operators to oversee the machinery. The "Intrusion detection systems (IDS)" have a significant challenge when it comes to detecting "cyber assaults on industrial control systems (ICSs)" because of the distinctive designs of these systems. It is important to note that the "Stuxnet attack" was the first "cyber-warfare weapon", making it stand out among the subsequent attacks on "Industrial Control Systems (ICS)". According to Nourian and Madnick (2018), the intention of the cyber attack known as Stuxnet was distinct from that of previous cyber attacks since it was most likely intended to impede Iran's nuclear programme. There are a number of different actors that might possibly launch attacks on "industrial control systems (ICSs)". These actors include nation-states, competitive businesses, internal criminals, and hacktivists. The "economy, national security,

and public health and safety" might all be adversely impacted if the "Industrial Control System (ICS)" were to be hacked. There have been "power outages, explosions, and chemical spills" that have resulted from compromised industrial control systems. If you want your business to be dependable, safe, and adaptable, secure ICSs are absolutely necessary. When it comes to protecting "industrial control systems (ICSs)" against assaults, it is very necessary to have "intrusion detection systems (IDS)" that are particularly developed for them. These systems should take into consideration the distinctive design of ICSs, as well as their real-time operation and the ever-changing environment.

DISCUSSION & CONCLUSION

It is becoming increasingly common for cybercriminals to target computer users by employing social engineering techniques and advanced technology tools. An ever-increasing number of cybercriminals also possess a high level of education and financial resources. It has been established that cybercriminals are capable of utilising sophisticated infrastructure, concealing their communication, and putting themselves in a distance from unlawful revenues. In order to ensure the safety of computer systems, it is necessary to implement "sophisticated Intrusion detection systems" that are able to "identify contemporary malware". It is vital to have a comprehensive understanding of the benefits and drawbacks of the current research being conducted on "Intrusion Detection Systems (IDS)" in order to construct and develop IDS systems that are successful. This article provides a complete analysis of the several approaches, categories, and technologies that are utilised in "Intrusion detection systems", including an analysis of the benefits and drawbacks associated with each of these techniques and technologies individually. In this article, various different machine learning approaches that have been proposed for spotting zero-day threats are investigated. These systems, on the other hand, can produce a number of false alarms or inaccurate detections, and they might have difficulty keeping up with the most recent information on developing threats. For the purpose of addressing challenges that are associated with IDS, we investigated previous research and investigated modern approaches to improve AIDS performance. For the purpose of "Intrusion Detection System (IDS) research", this article has conducted an analysis of the public datasets that are employed the most frequently. This analysis includes the data collecting methods, assessment outcomes, and limits of these datasets. The ongoing development of typical operations, which may become ineffective over time, necessitates the need for datasets that are both up to date and thorough, and that include a wide range of malware behaviours. The current machine learning algorithms are dependent

on training and analysing data from earlier datasets such as “DARPA/KDD99”, which do not cover modern “malware activities”. As a result, a new malware dataset is necessary because these approaches only use data from older datasets. In order to conduct testing, only datasets from 1999 that are accessible to the public are used. This is because there are no alternative possibilities that are considered appropriate. Even though these datasets were originally thought to be benchmarks, they do not reflect zero-day attacks that are occurring in the modern day. Despite the fact that it includes a number of new assaults, the ADFA dataset is insufficient. In light of this, it is possible that doing AIDS testing using these datasets may not result in an appropriate review, which may lead to inaccurate statements regarding the efficacy of the medications. In addition to this, the study analyses four typical escape strategies to establish how successful they are against the most recent “Intrusion Detection Systems (IDSs)”. It is essential for an effective “intrusion detection system (IDS)” to have a wide attack detection range and the ability to reliably identify invasions that make use of evasion strategies. One of the most significant challenges associated with this topic is the development of “intrusion detection systems (IDSs)” that are capable of successfully countering evasion strategies.

REFERENCES

1. Abbasi, J. Wetzels, W. Bokslag, E. Zambon, and S. Etalle, “On emulation-based network intrusion detection systems,” in Research in attacks, intrusions and defenses: 17th international symposium, RAID 2014, Gothenburg, Sweden, September 17–19, 2014. Proceedings, A. Stavrou, H. Bos, and G. Portokalidis, Eds. Cham: Springer International Publishing, 2014, pp. 384–404.
2. A. Aburomman and M. B. Ibne Reaz, “A novel SVM-kNN-PSO ensemble method for intrusion detection system,” Appl Soft Comput, vol. 38, pp. 360–372, 2016/01/01/ 2016.
3. Adebowale A, Idowu S, Amarachi AA (2013) Comparative study of selected data mining algorithms used for intrusion detection. International Journal of Soft Computing and Engineering (IJSCE) 3(3):237–241.
4. Agrawal S, Agrawal J (2015) Survey on anomaly detection using data mining techniques. Procedia Computer Science 60:708–713.
5. M. Ahmed, A. Naser Mahmood, and J. Hu, “A survey of network anomaly detection techniques,” J Netw Comput Appl, vol. 60, pp. 19–31, 1// 2016.

6. A. Alazab, J. Abawajy, M. Hobbs, R. Layton, and A. Khraisat, “Crime toolkits: the Productisation of cybercrime,” in 2013 12th IEEE international conference on trust, security and privacy in computing and communications, 2013, pp. 1626–1632.
7. Alazab, M. Hobbs, J. Abawajy, and M. Alazab, “Using feature selection for intrusion detection system,” in 2012 international symposium on communications and information technologies (ISCIT), 2012, pp. 296–301.
8. Alazab A, Hobbs M, Abawajy J, Khraisat A, Alazab M (2014) Using response action with intelligent intrusion detection and prevention system against web application malware. Information Management & Computer Security 22(5):431–449.
9. S. A. Aljawarneh, “Emerging challenges, security issues, and Technologies in Online Banking Systems,” Online Banking Security Measures and Data Protection, p. 90, 2016.
10. Annachhatre, T. H. Austin, and M. Stamp, “Hidden Markov models for malware classification,” Journal of Computer Virology and Hacking Techniques, vol. 11, no. 2, pp. 59–73, 2015/05/01 2015.
11. Bajaj K, Arora A (2013) Dimension reduction in intrusion detection features using discriminative machine learning approach. IJCSI International Journal of Computer Science Issues 10(4):324–328.
12. Bhuyan MH, Bhattacharyya DK, Kalita JK (2014) Network anomaly detection: methods, systems and tools. IEEE Communications Surveys & Tutorials 16(1):303–336.
13. Breiman L (1996) Bagging predictors. Machine Learning, journal article 24(2):123–140.
14. J. Camacho, A. Pérez-Villegas, P. García-Teodoro, and G. Maciá-Fernández, “PCA-based multivariate statistical network monitoring for anomaly detection,” Computers & Security, vol. 59, pp. 118–137, 6// 2016.
15. O. Can and O. K. Sahingoz, “A survey of intrusion detection systems in wireless sensor networks,” in 2015 6th international conference on modeling, simulation, and applied optimization (ICMSAO), 2015, pp. 1–6.
16. L. Chao, S. Wen, and C. Fong, “CANN: an intrusion detection system based on combining cluster centers and nearest neighbors,” Knowl-Based Syst, vol. 78, pp. 13–21, 4// 2015.

17. S. Chebrolu, A. Abraham, and J. P. Thomas, “Feature deduction and ensemble design of intrusion detection systems,” Computers & Security, vol. 24, no. 4, pp. 295–307, 6// 2005.
18. W.-H. Chen, S.-H. Hsu, and H.-P. Shen, “Application of SVM and ANN for intrusion detection,” Comput Oper Res, vol. 32, no. 10, pp. 2617–2634, 2005/10/01/ 2005.