Intrusion detection systems: A survey and taxonomy

By **Mohammed Tahmid Ahmed**

Abstract: With the exponential advancement in computer technology, people are now worried about revealing confidential data on the internet, making computer security an essential tool to detect these threats and using intrusion detection systems (IDS) to identify different types of malicious traffic. This paper will survey and propose the taxonomy to show modern IDS and their threats.

Keywords—IDS, intrusion, threats, signature-based, anomaly-based, deployment, engineer approach, malware

1. Introduction

The usage of networks around the world has increased significantly to where there are about 4.93 billion pages on the internet as indexed by the major engines [1] and made people change their lifestyles. It is well-established in education, finances, and other vital fields. However, several threats can affect network users, which can be considered dangerous. Therefore, network security has become critical to information sharing and human interaction. For that reason, the creation of the intrusion detection system would inspect outbound and inbound traffic for any malicious activity and detect the stunts before occurring [7].

Firstly, something is currently like the intrusion detection system (IDS), the intrusion prevention system (IPS). There is a clear distinction between IDS and IPS, as intrusion detection is the process of monitoring the network for any potential intrusion and an intrusion detection system purpose is to automate the intrusion detection [2]. The intrusion prevention system has all the abilities of an IDS and would try to prevent malicious traffic from entering the network. This paper will be focusing on surveying and classifying IDS techniques and then evaluating them.

Another worthy mention is that there are many types of intrusion, such as gaining unauthorised access to a device using the password, meaning that the intruder has proven his identity to the device as the actual user, and this would be called a masquerade [9]. Another intruder can be users that abuse their privileges and further damage the system.

2. Taxonomy of IDS

In the field of IDS, two main classified elements are detection methodology and deployment strategy. The detection methodology is where the system identifies the intrusion event and then alerts the intrusion detection system to secure the system, which can reduce the attack damage. The deployment strategy allows IDS to detect different threats by controlling different system areas [3], which can help IDS detect attacks at different system areas.

There are several classifications in IDS (as shown in figure 1), which have different methods of showing and using knowledge about intrusion detection. Signature-based systems need to store data about the attack patterns, and the Stateful protocol (specification-based) use the formal description of allowed usage [3]. An anomaly-based system requires an existing and proper profile behaviour and Rule-based use rule semantics, either fuzzy or crisp [3].

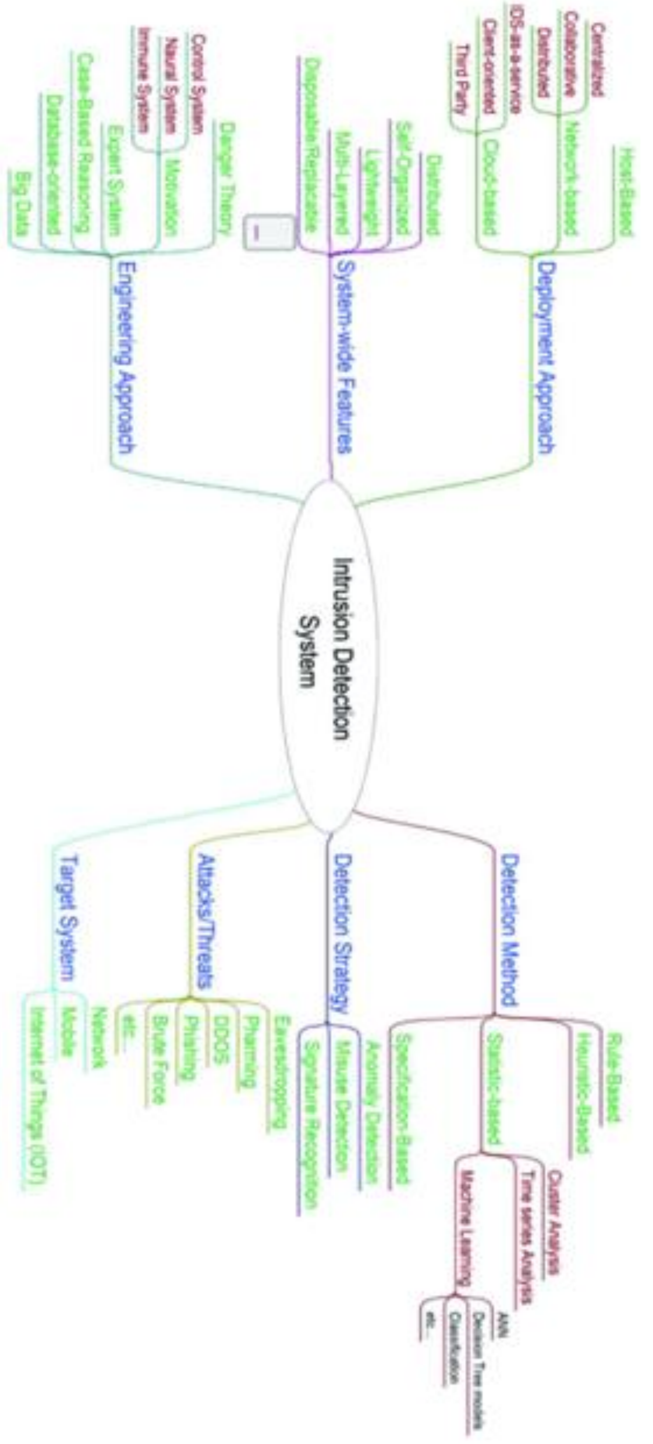

Figure 1: A taxonomy of IDS research (Updating the Taxonomy of Intrusion Detection Systems [3])

A good taxonomy includes the attack types and how the IDS would handle these types, as researching attacks can help the IDS improve against these attacks. Another classification will be the deployment approach, where the IDS handles specific systems to monitor or record the attack path [3]. Furthermore, the IDS will be different when the implementation technology choice depends on the system's best.

Figure 1 shows all types of elements included in the IDS. This paper will discuss several topics in detail. A. Phadke and S. Ustymenko have explained and created figure 1 based on considerable research and calculations to create an updated taxonomy of the IDS [3].

3. IDS METHODOLOGIES

As mentioned before, intrusion detection is a process that monitors the network or system and analyses them for any possible signs of malicious traffic or possible incident, which could be a violation of the computer security policy. These incidents can be malware (e.g., spyware, worms), and attackers will gain access to the private network and possibly damage or steal information from the network. However, some malicious traffic or incident can be just a user typing in an incorrect address and accidentally trying to connect to that system, and this can be seen as an incident, even though there is no malicious intent [4].

There are different types of Intrusion detection methodologies, but the most common ones are into three significant categories: signature-based detection, anomaly-based detection and stateful protocol analysis, as stated by Hung-Jen Liao et al. [2]. Their description is as follows: signature-based detection detects a pattern known to be an attack towards the system, and then the system would compare the pattern gained from the previous attacks to figure out whether it is an attack or a false alarm. Anomaly-based detection will detect irregular behaviours. If it is outside the expected behaviour, it will then compare it with other previous anomalies and the regular activities that happen, which would decide whether it is an attack. Stateful protocol analysis adds a stateful inspection that analyses protocols at the application layer and traces the protocol states.

3.1. SIGNATURE-BASED DETECTION

Signature-based detection is one of the most simple and effective methods against known attacks, and this is because it has a detailed history of the attack; this can reduce false positive alarms since the system knows the normal behaviour [8]. However, the downside of signature-based detection is that it cannot detect zero-day attacks when the database has no information on the threat variant [8]. Signature-based detection takes time to maintain a large amount of knowledge and would need to be updated regularly, according to Hung-Jen Liao et al. [2]. A potential solution for signature-based detection is anomaly-based detection, which depends on the user's behaviours if it is acceptable or not.

3.2. ANOMALY-BASED DETECTION

Anomaly-based detection, also known as behaviour-based, can detect new threats and it can detect privilege abuse. This detection is less dependent on the operating system, and if the system user is outside the set standard action, it will trigger a caution [3]. However, anomaly-based detection can be difficult because when rebuilding the behaviour profile, it would be available during that time. The profile accuracy cannot be weak and can occur by having an excessive number of changes or a considerable number of events [2]. This state that the profiles would need to be accurate based on the user expected actions. Another downfall of anomaly-based detection is that it will require extensive profile maintenance, which can be pretty slow and tiresome [3].

3.3. STATEFUL PROTOCOL

The stateful protocol will trace the protocol states and distinguish unexpected sequences of commands, which is helpful. On the other hand, this is consuming and unable to inspect attacks that go with the protocol behaviours. The stateful protocol might not be compatible with some operating systems [2].

3.4. DEPLOYMENT APPROACH

A host-based IDS can determine the targeted system and record the attacker's access path. However, there are downfalls of the host-based IDS, which would be that the system will have performance issues when the incomplete system model would require to make an accurate map of the system and then all the host layers will need to be performed on, which will take up system resources [3].

A network-based IDS can oversee the whole network, allowing an overview of the system, allowing identifying attacks more quickly. The issue of the network-based IDS is that the larger the data transmissions rate, the more complex the perfect placement will be [3].

4. ENGINEERING APPROACHES

4.1. EXPERT SYSTEM-BASED INTRUSION DETECTION SYSTEM

Expert system-based IDS uses a decision based on a security expert, which is done by gaining data that will help with the decisions to determine the point of intrusion. The system includes an interface, rule-based, inference engine and interpreter [3].

4.2. MULTI-LAYER IDS

Multi-layer IDS are where the system has multiple layers of defence; the first layer includes an application and component control and a process of termination [3]. The application and component control will monitor the behaviour of the application and will inspect the components of the application. If the application tries to execute the termination process will prevent it from doing so [3]. The second layer is the monitoring of hidden processes and control agents for the protection of physical memory stated by A. Phadke and S. Ustymenko [3]. The third level is a firewall and antivirus system that holds the signature-based database [3]. The final layer of the multi-layer IDS is the adaptive learning system, where the system learns the behaviour between self and non-self [3].

4.3. DATA-BASED IDS

Data-based IDS is records of the user or records of historical data used to analyse and understand the user's pattern, leading to the creation of the anomaly-based and misuse detection rule [3]. According to A. Phadke and S. Ustymenko, a suitable method of IDS database creation is to create a honeypot to lure in attackers and then analyse their behaviour, which improves the IDS against future attacks [3].

4.4. DANGER THEORY

The danger theory aims to reduce the number of false-positive reports. The danger theory prioritises the danger signals, which is based on the correlation of the signals from the theory. The system collects the data signals from the host and networks, which then correlates the alert signals. The system would minimise the actual negative rate by filtering the detectors, giving an accurate number of detectors too [3].

5. Intrusion detection evasion techniques

This section includes how cybercriminals avoid the IDS and gain access to the system or the network. The technique used by cybercriminals is flooding, fragmentation, encryption and obfuscation [8], created to avoid the IDS, allowing the malicious software to execute on the system and causing damage to the system or stealing sensitive data. When developing an IDS, one should understand the evasion techniques when creating or using techniques to detect these evasions.

5.1. Flooding

Flooding is where the attacker will overwhelm the IDS and cause it to malfunction allowing all traffic to get through. The most used methods can be spoofing the legitimate user datagram protocol (UDP) and the internet control message protocol (ICMP). The attacker can hide their malicious traffic during the flooding, making it hard for the IDS to detect the malicious traffic [8].

5.2. Fragmentation

Fragmentation is smaller packets that make up a packet, and then the fragmented packets are created again at the IP layer before the packet heads to the application layer. Attackers will use fragmentation to send malicious packets undetected when the fragmentation overwrites, overlaps or has a timeout. The attack replaces the information in the packet with a malicious packet [8].

5.3. Encryption

Encryption allows securing digital data, which allows having privacy. The attacker uses this method to escape detection and hide their attacks that will target a specific system [8]. Metke and Ekl stated that intrusion detection systems could not read attacks using HTTPS [11], and the IDS could not interpret the encrypted traffic.

5.4. Obfuscation

The obfuscation technique can avoid detection, which hides the code by making it challenging to analyse the message [12]. Obfuscation can change the code but can still do it objectively whilst changing, which can avoid the IDS. A signature-based IDS cannot detect obfuscation techniques as it requires the malware pattern in its database. However, the problem is that the obfuscation malware can change itself; thus, a new signature is created [8].

5.5. Challenges

Intrusion detection evasion techniques are still a considerable threat against intrusion detection systems as signature-based intrusion detection systems need to find the original signature or detect the new signature created by the malware [8]. To understand what happens after the evasion technique, we need to look deeper into computer attacks and understand how the IDS is affected.

6. Types of computer attacks

This section categorises the types of attacks based on the targets and activities of the attacker. There are several attack types classified into multiple classes: denial of service (DoS) attacks, phishing attacks, probing attacks, user-to-root (U2R) and remote-to-local attacks, brute force attacks and malware. These attacks can be passive, active or both. A passive attack is where the attacker has unauthorised access and can read the system data without taking the system resources. The active attack is where the attacker will use external components to gain access to the system; this would be worms, trojan and distributed denial of service (DDoS) [3,5]. The IDS must face these threats and expect how there usually are executed.

6.1. DoS attack

DoS attacks attempt to shut down networks services by restricting, blocking or flooding the network [3,5,8]. According to Ali Movaghar, there are two types of DoS attacks: networking attacks, which exploit existing network protocols, and operating system attacks, which targets bugs in a specific operating system [5]. An example of a networking attack is an SYN flood attack. The attacker will open multiple connections to that network, but these connections are half-open; this will lead to the system spending the resources on the half-opened link and then causing the system to be unresponsive. An example of an operating system attack is a teardrop, which is an attack that floods requests from the network and causes them to be unavailable; it does this by exploiting the vulnerability of TCP/IP fragmentation [5]. Ali Movaghar noted that it should focus on two main problems by analysing DoS attacks. The first main focus is early detection and being able to identify ongoing DoS activities. The second main focus is the response mechanism to reduce the effect of the DoS attack, achieved by blocking the packets linked to the attack, and it would slow down the attack [5].

6.2. Phishing attack

Phishing attack objectives are to steal sensitive data and then to be able to launch malicious software onto the system. There are different types of phishing: clone phishing, spear phishing, and whaling phishing. Clone phishing is where the user's address is a fake duplicate of the account. An example is when the user receives a legitimate email, the attacker will forge one and have malicious links [3]. Spear phishing is where the attacker spoof, targeting an individual or an organisation to access their sensitive information. The final type of phishing is whaling phishing that targets high-profile groups and encourages the user to commit an action where they send funds to the attacker [3].

6.3. Probing attack

When the attacker uses probing, its primary role is to gain information about the system or network, which is done by identifying the IP addresses and collecting data from them. The information given would allow the attacker to find vulnerabilities in the system [5]. Most intrusion detection systems can now detect known scanning tools, and attackers avoid this by lowering the transmission frequency, which would lower the chances of alerting the IDS [5].

6.4. U2R and R2L attack

U2R attacks target non-privileged users and gain access, then use malicious software to gain root access on a system, gaining user-level access [8]. There is another attack similar to U2R, and it is remote-to-local (R2L) attacks, which sends packets to the targeted system and then gain privileged access onto the system [8]. U2R and R2L can avoid IDS entirely due to the lack of information and accuracy on the threat.

6.5. Brute force attack

A brute force attack is a trial-and-error method of gaining the password of the device or system. However, the stronger the password, the longer it would take to figure the password out. There are different types of brute force attacks: search attack, dictionary attack, and rule-based search attacks [8]. The usual force method would not work with modern-day IDS, so the attackers came up with a stealthy way to avoid the IDS.

6.6. Malware attack

There are different types of malware: trojans, worms and viruses. Trojans are computer programs that function as legitimate programs but have hidden codes that allow unauthorised access or malware release [6]. Worms' can spread from system to system without any human interaction; it would usually spread through spam emails and interfere with systems files, disrupting the computer's performance [6]. Worms achieved this by using the automatic packets sending and receiving feature found on computers [5]. Viruses are an opposite of a worm as their replicate themselves by human interaction, and the cause will be running malicious files or opening email attachments [5].

7. Conclusion

Today's cybercriminals are becoming more complex by using computer attacks to social engineering whilst hiding their own identity and theft. Thus, making personal users, governments, and companies beware of their existence; therefore, IDS is crucial for detecting the threat. The IDS would then need to be updated to defend against the threats, and to achieve this is to figure out the system's limitations.

This paper discussed a survey and the taxonomy of the IDS to provide insights into the system, and the fields separated into different topics: IDS methodology, deployment approach, engineering approach, intrusion detection evasion techniques, and types of computer attacks with their advantages, disadvantages, description and how it will affect the IDS. However, when completing the survey, there was an issue when discussing the approaches, which was that updating and creating a new signature for IDS about new threats will have awful accuracy or report false alarms leading to intrusions, due to this it can make it challenging to observe actual attacks and to improve from it. To further develop the IDS is to clearly understand the taxonomy of the IDS and create an appropriate IDS to detect malicious activities.

This paper has assessed the IDS techniques and threats, proving that the IDS would need further research due to the zero-day vulnerability. The IDS should be updated frequently to detect new variants of the threats, and the administrator of the IDS should know the evasion techniques used by the intruders to prevent damage to the system and IDS. Using traditional signature-based techniques may be outdated as new techniques are adopted to detect modern-day threats and record the attacker's path, thus improving the IDS further ; this can be inadequate for the internet of things system because parts of the system can affect the IDS, which will need further configuration [10].

References

[1] M. de Kunder, “The size of the world wide web,” http://www.worldwidewebsize.com/, accessed 21/11/2021

[2] Hung-Jen Liao, Chun-Hung Richard Lin, Ying-Chih Lin, Kuang-Yuan Tung, Volume 36, Issue 1, 2013, “Intrusion detection system: A comprehensive review, Journal of Network and Computer Applications”

[3] I. Ghafir, J. Svoboda, V. Prenosil, “A Survey on Botnet Command and Control Traffic Detection,” International Journal of Advances in Computer Networks and Its Security (IJCNS), vol. 5(2), pp. 75-80, 2015.

[4] A. Phadke and S. Ustymenko, 2021, pp. 1085-1091"Updating the Taxonomy of Intrusion Detection Systems," 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC)

[5] I. Ghafir, V. Prenosil, M. Hammoudeh, F. J. Aparicio-Navarro, K. Rabie and A. Jabban, “Disguised Executable Files in Spear-Phishing Emails: Detecting the Point of Entry in Advanced Persistent Threat.” International Conference on Future Networks and Distributed Systems. Amman, Jordan, 2018.

[6] I. Ghafir, M. Husak and V. Prenosil, “A Survey on Intrusion Detection and Prevention Systems,” IEEE/UREL conference, Zvule, Czech Republic, pp. 10-14, 2014.

[7] Ali Movaghar, 2008, “Intrusion Detection: A Survey”, https://www.researchgate.net/publication/232623012_Intrusion_Detection_A_Survey

[8] Nicholas Engle, 2020, “Computer Crimes”

[9] Mokhtar Mohammadi, Tarik A. Rashid, Sarkhel H.Taher Karim, Adil Hussain Mohammed Aldalwie, Quan Thanh Tho, Moazam Bidaki, Amir Masoud Rahmani, Mehdi Hosseinzadeh, A comprehensive survey and taxonomy of the SVM-based intrusion detection systems, Journal of Network and Computer Applications, Volume 178, 2021

[10] I. Ghafir, J. Svoboda and V. Prenosil, “Tor-based malware and Tor connection detection,” International Conference on Frontiers of Communications, Networks and Applications. Kuala Lumpur, Malaysia, pp. 1-6, 2014.

[11] Khraisat, A., Gondal, I., Vamplew, P. et al. Survey of intrusion detection systems: techniques, datasets and challenges. Cybersecur 2, 20 (2019).

[12] I. Ghafir and V. Prenosil, “DNS query failure and algorithmically generated domain-flux detection,” International Conference on Frontiers of Communications, Networks and Applications. Kuala Lumpur, Malaysia, pp. 1-5, 2014.

[13] Stefan Axelsson, 2000, Intrusion Detection Systems: A Survey and Taxonomy

[14] Bruno Bogaz Zarpelão, Rodrigo Sanches Miani, Cláudio Toshio Kawakani, Sean Carlisto de Alvarenga, A survey of intrusion detection in Internet of Things, Journal of Network and Computer Applications, Volume 84, 2017, Pages 25-37

[15] I. Ghafir and V. Prenosil, “Advanced Persistent Threat Attack Detection: An Overview,” International Journal of Advances in Computer Networks and Its Security (IJCNS), vol. 4(4), pp. 50-54, 2014.

[16] A. R. Metke and R. L. Ekl, "Security Technology for Smart Grid Networks," in IEEE Transactions on Smart Grid, vol. 1, no. 1, pp. 99-107, June 2010

[17] I. Ghafir and V. Prenosil, “Blacklist-based Malicious IP Traffic Detection,” Global Conference on Communication Technologies (GCCT). Thuckalay, India: pp. 229-233, 2015.

[18] Kim, Danny & Majlesi-Kupaei, Amir & Roy, Julien & Anand, Kapil & Elwazeer, Khaled & Buettner, Daniel & Barua, Rajeev. (2017). DynODet: Detecting Dynamic Obfuscation in Malware. 97-118.

[19] Mohammad Masdari, Hemn Khezri, A survey and taxonomy of the fuzzy signature-based Intrusion Detection Systems, Applied Soft Computing, Volume 92, 2020

[20] D. Jin, Y. Lu, J. Qin, Z. Cheng and Z. Mao, "KC-IDS : Multi-layer Intrusion Detection System," 2020 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS), 2020, pp. 1-5

[21] S. Eltanani and I. Ghafir, "Aerial Wireless Networks: Proposed Solution for Coverage Optimisation," IEEE Conference on Computer Communications Workshops”, IEEE, 2021.

[22] I. Ghafir and V. Prenosil, “Advanced Persistent Threat and Spear Phishing Emails.” International Conference Distance Learning, Simulation and Communication. Brno, Czech Republic, pp. 34-41, 2015.

[23] M. Hammoudeh, I. Ghafir, A.Bounceur and T. Rawlinson, “Continuous Monitoring in Mission-Critical Applications Using the Internet of Things and Blockchain,” International Conference on Future Networks and Distributed Systems. Paris, France, 2019.

[24] I. Ghafir and V. Prenosil, “DNS traffic analysis for malicious domains detection,” International Conference on Signal Processing and Integrated networks. Noida, India: pp. 613 - 618, 2015.

[25] U. Raza, J. Lomax, I. Ghafir, R. Kharel and B. Whiteside, “An IoT and Business Processes Based Approach for the Monitoring and Control of High Value-Added Manufacturing Processes,” International Conference on Future Networks and Distributed Systems. Cambridge, United Kingdom, 2017.