# THE OVERVIEW OF INTRUSION DETECTION SYSTEM METHODS AND TECHNIQUES

Ali Elsadai[1, *],
Juma Ibrahim[2],
Fathi Hajjaj[2],
Petar Jakić[1]

[1]Singidunum University,
Belgrade, Serbia
[2]School of Electrical Engineering,
Belgrade University,
Belgrade, Serbia

Abstract:
System and network security is still the main concern for researchers and system administrators. Different types of software and hardware devices are used to eliminate the dangers of hacking, attacks, breakthroughs, and cybercrimes. With the increasing creativity of attackers, the development of a complete security system is impossible. In network security, no other tools are available as Intrusion Detection System (IDS), which has the ability to locate and identify malicious activities by examining network traffic. It provides administrators with reliable network monitoring and control. In this paper, we give an overview of the Intrusion Detection System principles, types, methods, and detection techniques.

Keywords:
Flow, IDS, IPS, Signature-based, Anomaly-based.

## 1. INTRODUCTION

In light of the tremendous progress in technology and communications, most of our daily activities have become dependent on the Internet and its applications. It has entered in everything and almost there is no home free of this technology to benefit from its services in everything related to daily life such as news, Sport, art, education, media exchange and makes free phone calls with other people and groups of people from all over the world. The use of the Internet saves a lot of effort and time where you can communicate with other people, transfer files and access different services from anywhere, at any time where you have access to the Internet.

With this amazing progress in technology and communications and with these services and possibilities provided by the Internet, in contrast, there are many problems and obstacles that affect the utilization of the Internet service. For example, the disruption of a network device such as router or switch and loss of communication with these devices or other devices likes hosts or servers. Such problems are predictable by the network team, and they are ready to overcome such problems or provide alternative solutions. There is another kind of invisible problems that come to its utterance without warning and cannot be discovered by ordinary users. Also, specialists difficultly predict and identify the so-called

Correspondence:
Ali Elsadai

e-mail:
aelsadai@singidunum.ac.rs

DOI: 10.15308/Sinteza-2019-155-161

cyber-attacks or hackers who target computer networks to stop the services and servers of these networks or the information that is transmitted through the network to prevent possible sabotage and corruption.

This conflict between good and evil doesn't stop until the end of the cessation of life. Because at the beginning of each new day, systems and networks face a new type of these electronic attacks in an attempt to break into a database and steal data or sabotage or stop services and web sites.

But as we mentioned earlier, in this amazing progress in technology and communications, the high-speed of the Internet networks and the huge amount of data passing through computer networks, require the presence of reliable techniques for monitoring and controlling the network where the specifications of speed and accuracy are not affected by the flow of data through the network and does not affect users experience concerning speed and quality.

Many of the scientific papers published in this regard [1][2][3][4][5][6], searching for a solution to protect computer networks or, more precisely, how to distinguish between normal traffic from anomaly traffic passing through the network. Traditional security tools such as firewalls and anti-virus programs are no longer capable of detecting these sophisticated attacks. Intrusion Detection Systems (IDS) is imperative to achieve confidentiality, integrity and availability of data and services. So, it's considered as an important component for computer networks; it is a kind of protection for network systems, and devices and a set of measures to be taken when an attempt is made to penetrate the security objectives of a network.

## 2. FLOW MONITORING

The concept of network flows was introduced by Cisco and is currently standardized by the Internet Engineering Task Force (IETF). According to the IETF IPFIX working group, "A flow is defined as a set of IP packets passing an observation point in the network during a certain time interval" [7]. All packets belonging to a particular flow have a set of common properties. In the simplest form, these properties are the source and destination addresses and ports. Flow is typically defined as a unidirectional sequence of packets, which means that there are two flows for each connection between two endpoints – one from the server to client and one from the client to the server.

Recently, bidirectional flows (one record for each session between two endpoints) are also supported by vendors.

There are two approaches to network traffic monitoring, namely, packet-based and flow-based.

1. Packet-based is the traditional representation way of network traffic which monitors the whole packets' headers and payloads in the form of time-stamped row packets. All network packets passing through a particular monitoring device are captured without losing or filtering any information. Packets are captured in the same format that occurs in the wire containing information of layer 2 to layer 7 of the Open Systems Interconnection (OSI) model. Therefore, capturing should be made at an edge point of the network where incoming and outgoing traffic can be monitored [8].
2. TcpDump [9] and Wireshark [10] are two common software applications that can be installed on the capturing point for the packet-based capturing. The main advantage of the packet-based representations is the availability of all necessary information for IDS such as payload, header, and time of the packet.
3. Flow-based traffic representation is a more advanced representation way of network traffic and depends on formatting the traffic in flow forms. A flow is defined as a set of packets that shares the same characteristics. These characteristics vary according to the targeted attacks and the IDS detection criteria. For example, the IPv4 dataset KDD99 [11] uses five-tuple keys to form its records (flows) including IP source, IP destination, port source, port destination, and protocol values as the flow characteristics. Unlike packet-based, flow-based representation provides only a chosen set of information to achieve the detection of the targeted intrusion [12]. The main advantage of flow-based representation is that it does not form a lot of traffic compared to packet-based representation.

The flow-based network monitoring is more scalable in the context of network speed. This approach is based on the ability of network devices to aggregate packets in flows. Each flow is cached by a device and when it is finished or timeout is exceeded, it is exported to an element called the collector. Table-1 compares the two traffic representations.

| Flow-based NIDS | Packet-based NIDS |
|---|---|
| **Pros** | **Pros** |
| NIDS operability is not influenced by encrypted payloads | Consists of complete payload and header up to layer 7 |
| Less privacy problems | Data available to NIDSs without delay |
| Scalability to high-speed network | |
| **Cons** | **Cons** |
| Limited data availability | Processing packets before any filtering accomplished |
| Additional component for pre-processing flows | Exposing confidential |
| Latency between capturing and the availability to the NIDS | No signature matching possible for encrypted payload cases |

## 3. WHAT IS AN IDS

Intrusion Detection is the process of monitoring events in a computer or network system and analysing them to identify any indications of potential violations, breaches, threats or imminent threats to computer security policies, acceptable use policies or security practices. Malicious software (such as worms and spyware), attackers gain unauthorized access to the system through the Internet, or users who are not authorized or authorized by the authorities trying to obtain privileges that are not their right. Although all of these incidents are inherently harmful, some other incidents may result in a coincidence, figure-1 shows the concept of IDS.

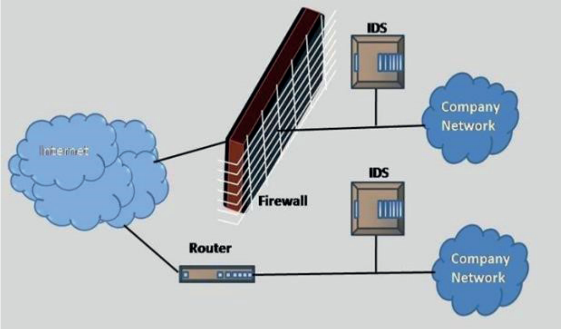

Figure-1 Intrusion Detection System

Intrusion Detection System (IDS) is a program that performs a monitoring operation. This process is to scan the data and compare it to a large database that contains logical rules or fingerprints for previous attacks. It checks the malicious traffic, and if it finds a match, sends a warning to the network administrator. The system does not deter or prevent attacks, but its only function is to detect, record and warn when it occurs.

Intrusion Prevention System (IPS) is a program that has all the capabilities of the Intrusion Detection System; it also has the ability to prevent intrusion. It has the ability to stop or block network traffic based on logic or fingerprint it is capable of detecting network traffic, analysing ports, recording events, checking and matching content, and can detect physical operations, attacks, and port scans all in real time and can be combined with other reporting and performance analysis equipment and records [13].

## 4. USES OF IDS TECHNOLOGIES

The main functions of the intrusion detection systems are the possibility of identifying possible incidents. For example, hackers can be detected if they have successfully penetrated the system because of security vulnerability. They report security to the system administrators quickly and do what is appropriate to minimize the damage caused by this intrusion. It also records the information so that it can be used by incident handlers later in terms of digital forensics; IDS settings can also be set to identify violations of the security policies of the institution. It can also monitor file transfers and identify those that may be suspicious, such as copying a large database to a user's mobile device. It can also challenge reconnaissance activities, which may indicate an imminent attack because some of the attacking tools and malicious programs, such as scanning activities of the host and ports to determine the targets for subsequent attacks. IDS may be able to block the survey and notify security officials, who can take action if necessary to change other security controls to prevent related incidents. Because reconnaissance activity is very frequent on the Internet, the survey is mostly conducted on protected internal networks.

In addition to identifying incidents and speeding up the response to these incidents, organizations found that other uses of IDS beside intrusion detection or system penetration are as follows:

Identifying security policy issues, IDS can provide a certain degree of quality control to implement a security policy, such as repeating firewall and alert databases when it sees network traffic that should have been blocked by the firewall but is caused by a firewall configuration error.

Documenting the Organization's Current Threat IDS records information about the threats it detects. Understanding the frequency and characteristics of attacks against computer resources in the organization is useful in identifying appropriate security measures to protect network resources. Information can also be used to educate management about threats faced by the organization.

Deter individuals from violating security policies; if individuals notice that their actions area unit is monitored by IDS techniques for security policy violations, they will be less probably to commit such violations because of the risk of disclosure.

Due to the increasing reliance on information systems and the proliferation of potential impacts of interventions against these systems and their attendant, IDS has become a necessary addition to the security infrastructure of all its institutions.

## 5. TYPES OF IDS

There are two main types of intrusion detection systems according to the place where it's located in case of Hardware or deployed if it is software, as shown in figure-2.

1. Host-based Intrusion Detection System (HIDS): monitors characteristics of a single computer, it is special software on a computer try to detect intruders; it monitors what happens in that computer with the aim of detecting suspicious activities. Generally, monitoring what happens in computers such as login, sessions, activities, transferring documents, file access and programs access.

   The main purpose of HIDS is usually only to monitor changes to basic operating system files. They may not have awareness of what applications have been installed or proprietary data to save, defining critical data on the assets and create a policy to detect changes in that data also may be using custom application.
2. Distributed host-based Intrusion Detection system (DHIDS): intrusion detection software is distributed in many computers which trying to monitor what happens and they report back to some other central computer to improve the chance of detecting the intrusions, it's a Host-based IDS on multiple computers with an organization LAN or internetwork.
3. Network-based Intrusion Detection system (NIDS): Monitor network traffic to identify suspicious activity, monitoring the packet through a network what is being sent and received, different points across the network monitoring the traffic and doing that with some sensors which are devices that basically keep track of the packets coming in or out, then some analysis of that traffic is taking place to identify whether it is an intruder or not by looking at the packet behaviour based on the traffic. The analysis may be done in the sensors or by sending information back to the management server that does the analysis for all of them.
4. Distributed Network-based Intrusion Detection system (DNIDS): Although the traditional design captures the packets, it is vulnerable to lose. In Distributed Design / Distributed Technology, it puts sensors in all network devices, communicating sensors exchange information with each other to report events and breakthroughs. A command console will be used as a centre for decision making and alarms; the difference between Distributed Network-based and Distributed Host-based is the place where the sensors/agents are located. If the IDS works on each device in the network and analysis the events of the operating system on which it works, then it is Host-based, If the IDS works on each device in the network and analysis all the events it receives from the network card whether it belongs to the operating system on which it works or not, then it is a Network-based.

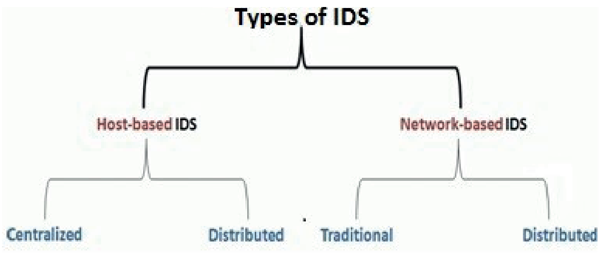

Figure-2 Types of IDS

## 6. OVERVIEW OF THE IDS DETECTION METHODS

Network traffic measurement is a well-developed and highly active field for network security specialists. Firewalls which were considered the main line of defence in networks are no longer able to carry out this mission as required. With the variety of methods and processes of cyber-attacks, also the intelligence put on the firewall does not provide protection from internal network usage. The intrusion detection system which is one of the measurement tools used for measuring network traffic operates through two main approaches:

1. Signature-based Intrusion Detection System (sometimes called a knowledge-based or misuse-detection) in this type, the information that is already present is used to represent specific and known types of existing attack patterns (signatures). The data moved through the network are compared with these patterns or signatures, if there is a consensus, it means that it is malicious data and the instructions are issued to stop it. One of the features of this type is the accuracy in identifying attacks but it suffers from a lack of detection of new attacks that are not stored patterns or signatures, which require an update of signatures and patterns of new attacks.
2. Anomaly-based Intrusion Detection System (or Behaviour-based), this type works by creating a profile from normal patterns and considers any deviation from this profile as an attack. Advantages of this system, it has the ability to detect new attacks but suffers from the problem of false positive alarm.

Some researches propose hybrid intrusion detection approach. Most of them focus on the anomaly detection approach. It is better than the signature-based detection [14].

In recent years different approaches of IDS based on anomaly detection have emerged using data mining, entropy, machine learning, statistical analysis, and artificial intelligence techniques such as genetic algorithms, artificial intelligence, etc. Table-2 shows the comparison between these two approaches of intrusion detection systems.

Since most of the current existing intrusion detection systems in enterprises are signature-based, they are still unable to detect new forms of attacks even if these attacks are derived slightly from the known attacks. Furthermore, the signature-based intrusion detection methods are inefficient in detecting cryptographic traffic and zero-day attacks, so they are unable to detect modern and recent attacks and need to be updated continuously. It is reactive in the sense that they use a set of signatures and needs to be updated frequently.

## 7. DATASETS

This section provides an overview of publicly available, network-based evaluation data sets and traffic generators, most of the researches use a synthetic dataset to conduct research and test their studies. The synthetic dataset is collected for this purpose and certainly will not be given the accuracy required as the real data to detect the attacks, but the real data is not easy to obtain for the most important reasons is the security reasons, there are many datasets but the famous and most widely used are:

| Signature-Based IDS (SIDS) | Anomaly-Based IDS (AIDS) |
|---|---|
| It depends on a specific pattern of the behaviours in the effort to distinguish between legitimate and malicious network traffic. | The main concept of AIDS is that attackers should generate network activities that are different from normal user's activities. |
| **Pros and Cons:** | **Pros and Cons:** |
| 1. It is inefficient to detect attacks because of their detection nature that is limited by attacks that are previously recorded in their database. | 1. Detecting new (zero-day) attacks, if their activities fall out of the normal predefined profile without any need for signature matching. |
| 2. Unable to detect zero-day (new) attacks which do not have a stored pattern in the database. | 2. More overhead. |
| 3. Very accurate. | 3. Processing capacity. |
| 4. Very fast. | 4. May not detect minor changes to system variables. |
| 5. Easy to configure. | 5. Generate many false positives alarms. |

1. KDD 99 was created in 2009; it represents the recording of the state of communication between two devices for a limited period of time using one of the known protocols and it is characterized by information known as attributes [11]. This dataset suffers from many flaws. The most important ones are outdated and do not contain any contemporary and attack behaviours. This explains the high accuracy rate achieved by many works.
2. UNSW-NB15 This Dataset is newer than the KDD cup 99, it was prepared in 2005 and contains 9 modern attacks compared to 14 attacks in KDD cup 99 and contains 2540044 records and 49 attributes [15].
3. In 2009 the first flow-based data set was introduced [12]. The authors collected network data; this data set is based on data collected from a real honeypot (monitored trap) featuring HTTP, SSH and FTP services. Some authors have gathered about 14 million malicious flows however most of them cited to the activity of internet and network scanners
4. CTU-13 malware is another flow-based data set which contains normal and malicious network traffic for the modern attacks like Denial Of Service (DOS), Distributed Denial Of Service (DDOS) and Robot Networks (BOTNET's), this dataset which has just been made public is much richer and consist of traces of several bots, namely Neris, Rbot, Sogou, Murlo, Menti [16].

## 8. CONCLUSION

Modern research, therefore, focuses on the development of new techniques, algorithms and security tools that uses intelligent methods such as network anomaly detection based on behaviour analysis which recognized as a mandatory part of modern security protection. Anomaly detection is one of the data mining tasks which analysis the large quantities of data to identify items, events or observations which do not conform to an expected pattern. A large number of scientific papers address this subject and reports that entropy and machine learning approaches provide promising solutions to this problem to overcome the weakness of traditional volume and rule-based approaches to network flows analysis

The main challenges with the anomaly-based intrusion detection system are setting up precise boundaries between normal and anomaly network behaviour, also to avoid false positive errors and the low rate for detecting the anomalies. Another problem with this technique the computation time needs for anomaly details extraction and root-cause identification.

## REFERENCE

[1] P. Barford, J. Kline, D. Plonka, and A. Ron, "IN PROCEEDINGS OF ACM SIGCOMM INTERNET MEASUREMENT WORKSHOP 2002 1 A Signal Analysis of Network Traffic Anomalies," pp. 1–12, 2002.

[2] T. Pevny, M. Rehak, and M. Grill, "Identifying suspicious users in corporate networks," *Proc. Work. Inf. forensics Secur.*, pp. 1–6, 2012.

[3] M. A. Ambusaidi, X. He, P. Nanda, and Z. Tan, "Building an intrusion detection system using a filter-based feature selection algorithm," *IEEE Trans. Comput.*, vol. 65, no. 10, pp. 2986–2998, 2016.

[4] L. Jadhav and C. M. Gaikwad, "Implementation of Intrusion Detection System using GA," *Int. Res. J. Comput. Sci.*, vol. 1, no. 1, pp. 1–5, 2014.

[5] J. Ibrahim, "SDN-Based Intrusion Detection System Literature review," *Infoteh-Jahorina*, vol. 16, no. March, pp. 621–624, 2017.

[6] M. A. Aydin, A. H. Zaim, and K. G. Ceylan, "A hybrid intrusion detection system design for computer network security," *Comput. Electr. Eng.*, vol. 35, no. 3, pp. 517–526, 2009.

[7] P. Bereziński *et al.*, "Network Anomaly Detection Using Parameterized Entropy To cite this version : HAL Id : hal-01405630 Entropy," 2016.

[8] O. E. Elejla, M. Anbar, B. Belaton, and B. O. Alijla, "Flow-Based IDS for ICMPv6-Based DDoS Attacks Detection," *Arab. J. Sci. Eng.*, vol. 43, no. 12, pp. 7757–7775, 2018.

[9] V. Jacobson, C. Leres, S. McCanne,: Tcpdump &LiBCAP. http://www.tcpdump.org. Accessed 03 May 2016

[10] L. Chappell, G. Combs, Wireshark network analysis,. https://www.wireshark.org. Accessed 22 June 2017

[11] KDD, C.; Nsl-Kdd.: Nsl-Kdd: Dataset for network–based intrusion detection systems. (1999). http://www.unb.ca/cic/research/datasets/nsl.html. Accessed 30 Mar 2017

[12] A. Sperotto, Flow-based intrusion detection. Ph.D. dissertation, Centre for Telematics and Information Technology, University of Twente, Netherlands (2010)

[13] M. Mohammed and A.-S. Pathan, "Intrusion Detection and Prevention Systems (IDPSs)," *Autom. Def. Against Zero-day Polymorphic Worms Commun. Networks*, pp. 47–84, 2013.

[14] P. Bereziński, B. Jasiul, and M. Szpyrka, "An entropy-based network anomaly detection method," *Entropy*, vol. 17, no. 4, pp. 2367–2408, 2015.

[15] M. Nour, J. Slay. "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)." *Military Communications and Information Systems Conference (MilCIS)*, 2015. IEEE, 2015.

[16] S. Garcia, M. Grill, J. Stiborek and A. Zunino "An empirical comparison of botnet detection methods", Computers and Security Journal, Elsevier. 2014. Vol 45, pp 100-123. http://dx.doi.org/10.1016/j.cose.2014.05.011