An overview of Intrusion Detection within an Information System: The Improvement by Process Mining

Nkondock Mi Bahanag Nicolas1 & Atsa Etoundi Roger1

1Department of Computer Science, University of Yaounde I, Cameroon

Correspondence: Nkondock Mi Bahanag Nicolas, Department of Computer Science, University of Yaounde I, Cameroon. E-mail: nicolas.nkondock@uy1.uninet.cm

Received: June 10, 2022   Accepted: July 25, 2022   Online Published: July 31, 2022
doi:10.5539/nct.v7n1p55   URL: https://doi.org/10.5539/nct.v7n1p55

Abstract

Information Systems handle big amount of data within enterprises by offering the possibility to collect, treat, keep and make information available. To realize these tasks, it is important to secure data from intrusions that can affect confidentiality, availability and integrity of information. Unfortunately, with the time, technologies are more used and various types of attacks act on it to create intrusion or misuses within Information Systems. Research in intrusion detection field is still looking for solutions of such relevant problems. The purpose of this paper is to present an overview of existing intrusion detection techniques compared to a new issue based on process mining used for event logs analysis to detect abnormal events that occurs on the system. events are classified accordingly to security policy established with fuzzy logic to build a set of fuzzy rules, for the definition of normal and abnormal events and then reduce the high level of false alerts.

Keywords: intrusion detection, fuzzy logic, false alerts, process mining, security policy

1. Introduction

Nowadays, enterprises use different technologies to enhance their business processes, by boosting the quality of service, in order to be more competitive on the market where needs of users or customers are permanently growing. Nevertheless, like a law of the nature, advantages usually generate some problems. In this case, while the quality of service is improved by the using of powerful technologies, security of data manipulated within information systems appears like a big challenge. All the actions that affect confidentiality, integrity and availability of information are considered as intrusions. Confidentiality concerns rights and authorizations of users while integrity is about the reliability of information and of course, every time, data must be accessible in real time by authorized users, this is availability. Every day information systems are the cible of several IT attacks. It can be made by internal or external attackers. All actions that are not authorized are considered as intrusive and naturally lead to a loss of quality of service. Moreover, wars in the world are actually managed mainly thanks to IT systems. We remember for instance an intrusion realized by the virus called Stuxnet in the Iranian nuclear program. It has affected that program during 2 years and has imposed big financial damage. Another example of intrusion is the one realized by the American Edward Snowden in NSA system. His action was considered as intrusive because, he has performed some actions like accessing to certain sensible data without having permissions. Similar situations are legion in the world, appear every day, every hour and generate several bad effects within organizations. These situations show that intrusion detection is still a big issue in Information System management. A big amount of papers exists on the topic of intrusion detection, each of them using a specific approach and presenting some advantages and certain limits. The aim of this paper is to provide an overview of intrusion detection techniques used and how it can be improved on the basis of process mining (Van der Aalst, 2011) and fuzzy logic (Atre & Singh, 2016). This paper is organized as follows. Sections 2 and 3 respectively describe intrusion detection techniques and systems. Section 4 shows big challenges addressed by intrusion detection, while section 5 justify and presents the use of process mining for intrusion detection and how fuzzy logic is used to define security policy. The last section concludes the paper.

2. Intrusion Detection Techniques

2.1 Signature-Based Intrusion Detection

(Roger et al., 2013) The signature-based detection technique also known as misuse detection technique is used in detecting and catching intrusions in terms of the characteristics of known attacks or system vulnerabilities. Therefore, any action conform to the pattern of a known attack or vulnerability is considered as intrusive. This
technique refers to techniques that use patterns of known intrusions or weak spots of a system to match and identify intrusions. The sequence of attack actions or activities, the conditions that compromise an information system security, as well as the damage left behind by intrusions can be represented by a number of general pattern matching models. Signature based intrusion detection uses regular expressions. Known signature are stored in a database and every event is compared to this base. If there is a matching, then an alert is generated. The attack must be known to be detected. But, the rate of false positive is low. The power of an Intrusion Detection System using this technique depends on the management of the database containing signatures, it is not the case with Anomaly-based intrusion detection.

2.2 Anomaly-Based Intrusion Detection

(Roger et al., 2013) Anomaly detection is based on the normal behavior of an actor within an information system, for this end any action that significantly deviates from the normal behavior is considered like intrusion. The proposed approach is focused on a formal and sound description of resources that participate in the execution of identified activities. The anomaly based intrusion detection techniques allows to detect unusual behavior and thus have the ability to detect symptoms of attacks without specific knowledge of details. They also help in producing information that can in term be used to define signatures for misuse detectors. However, these techniques usually produce a large number of false alarms (Roger et al., 2013)(Kruegel, Valeur, & Vigna, 2004) due to the unpredictable behaviors of users and networks; moreover, they often require extensive training sets of system event records in order to characterize normal and abnormal behavior patterns.

2.3 Hybrid-Based Intrusion Detection

It is the association of misuse and anomaly detection techniques, like in (Yoo et al., 2014). The goal is to detect unknown intrusion by analysing and catching abnormal behavior occurring in the system without generating a big amount of false alerts, and at the same time, detect and catch intrusive activities by analyzing and verifying if their signatures are present in the database of the system, containing the list of signatures that represents non-authorized activities. Mainly, the aim of hybrid-based intrusion detection technique is to combine advantages of the precedent techniques and substantially reduce their limits. Research continues in this field and, actually, there is not a perfect hybrid-based intrusion detection model.

3. Intrusion Detection Systems

Intrusion Detection within an Information System consists on the monitoring of different events that occur in the considered system (Marinova-Boncheva, 2007). Intrusive events are the ones that contains irregular informations, in order not in conformity with security policy established on the base of the organizational strategy. Steps of intrusion detection process are (Marinova-Boncheva, 2007) a) Monitoring and analyzing traffic; b) Identifying abnormal activities; c) Assessing severity and raising alarm. These steps are executed permanently by the intrusion detection system in a cyclic way. It exists two main types of intrusion detection systems HIDS and NIDS.

3.1 HIDS

HIDS: Host-based Intrusion Detection System - Controls activities of a single equipment, like a computer. It helps the monitoring of abnormal activities occurring in a specific machine used in the information system. It can be a server or the computer of an administrator for instance. The task here is more axed on activities of different users. The interest of HIDS can be oriented on the monitoring of the operating systems or the applications.

3.2 NIDS

NIDS: Network-based Intrusion Detection Systems - Analyzes traffic existing between computers present in a network. It can detect irregularities like surcharge of the network or bad information present in transferred data.

3.3 HONEYPOT

It is a computer connected to the network implementing willingly a low level of security. The goal is to distract attackers in order to protect more sensitive computers. Moreover, a honeypot is a good way of discovering new techniques of attack and new tools.

4. Big Challenges of Intrusion Detection

Many kind of attacks can occur in an information system, and thus are considered as intrusive. These attacks disturb confidentiality, availability and integrity of information which are core characteristics of a secure system.

1) Confidentiality: Different users, resources in general have not the same access to data. Some of them are not intended to be known publicly. Many level of rights are implemented within an information system. Nevertheless, an attack can make accessible sensitive information to an unauthorized person or resource.
2) Availability: Every time, data, resources, the whole system must be accessible and able to produce good results.

3) Integrity: It implies that data are not modified or destroyed by an unauthorized action.

4) Authentication: Different users of the system must provide personal codes, proving their identity. The aim of authentication is to check provided information in order to be sure that the user of the system is allowed to have a certain view on it.

5) Non-repudiation: It is always possible to know the authors of all the actions realized within the Enterprise. The main categories of attacks founded in the literature are: DoS, DDoS, Scan, U2R (Marinova-Boncheva, 2007), Probe, zero-day attacks

4.1 DoS and DDoS Attacks
The goal of Denial of service (DoS) is to attack a system and saturate its resources such that, the considered system become unavailable. DDoS refers to Distributed Denial of Service.

4.2 U2R: User to Root Attacks
For this type of attack, the attacker has an access to a normal user account on the system and exploits vulnerabilities to illegally have a root access to the system. With a root access, the attacker can create several damages in the system and then disturb confidentiality, integrity or availability of data in the system.

4.3 R2L: Remote to Local attacks
In the class R2L attack, the attacker sends packets to a machine via the network to illegally get a local access. Thus, a remote machine considered as the attacker, has the possibility to send packets via the network and takes advantage of some weaknesses of another machine to gain access to a local account on that machine.

4.4 Probe Attacks
In a probing attack, the attacker scans a network to amass information suitable (to exploit vulnerabilities).

4.5 Zero-Day Attacks
Such attacks are discovered when it appears, because the system has not knowledge about it.

From the above-mentioned categories of attacks, the following table shows different types of specific attacks founded in the literature and that can be detected by a network or a host based Intrusion Detection System (Proctor, 2000).

<table>
<thead>
<tr>
<th>No.</th>
<th>Challenge</th>
<th>Type of Detection</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Unauthorized Access</td>
<td>NIDS</td>
</tr>
<tr>
<td>2</td>
<td>Unauthorized Login</td>
<td>NIDS</td>
</tr>
<tr>
<td>3</td>
<td>Jump-Off Point for Other attacks</td>
<td>NIDS</td>
</tr>
<tr>
<td>4</td>
<td>Data/Resource Theft</td>
<td>NIDS</td>
</tr>
<tr>
<td>5</td>
<td>Password Downloads</td>
<td>NIDS</td>
</tr>
<tr>
<td>6</td>
<td>Bandwidth Theft</td>
<td>NIDS</td>
</tr>
<tr>
<td>7</td>
<td>Denial of Service</td>
<td>NIDS</td>
</tr>
<tr>
<td>8</td>
<td>Malformed Packets</td>
<td>NIDS</td>
</tr>
<tr>
<td>9</td>
<td>Packet Flooding</td>
<td>NIDS</td>
</tr>
<tr>
<td>10</td>
<td>Distributed Denial of Service</td>
<td>NIDS</td>
</tr>
<tr>
<td>11</td>
<td>Abuse of Privilege Attack</td>
<td>HIDS</td>
</tr>
<tr>
<td>12</td>
<td>Contractors with Elevated Privileges</td>
<td>HIDS</td>
</tr>
<tr>
<td>13</td>
<td>Ex-Employee using Old Account</td>
<td>HIDS</td>
</tr>
<tr>
<td>14</td>
<td>Administrator Creates Back-door Accounts</td>
<td>HIDS</td>
</tr>
<tr>
<td>15</td>
<td>Inadvertent Privileges Granted</td>
<td>HIDS</td>
</tr>
</tbody>
</table>
The Critical Data access and Modification problem can be addressed by process mining.

5. Process Mining

5.1 The goal of Process Mining

The purpose of Process Mining is to provide methods, and tools to analyze data related to process execution. Such data are stored in structured files called event logs, that keep the trace of different events associated to process execution. It is the reason why, the father of Process Mining, Will Van der Aalst, wrote in (Jagadeesh Chandra Bose & Van der Aalst, 2009) that, event logs constitute the starting point of Process Mining. More, The Handbook of process mining (Van der Aalst, 2011) clearly defines its purpose which is mainly about process discovery, conformance checking and process enhancement. Process discovery is about the way to build a process model based on event logs. Once the model is built, it is compared to a normative model, that describes how the model was supposed to be, this action is realized according to conformance checking concept. The third use of process mining is process enhancement that provides solutions to improve business processes. With a security challenge, the conformance checking is about security issues, such that, every intrusion is a violation of the security policy that constitutes the normative model used to parse or check the descriptive one.

5.1 Fuzzy Logic for Security Policy

A set of rules is used to define Security Policy. The trace of task execution is kept in event logs to improve analysis. Then, fuzzy logic is used to categorize data and events occurring in the system (Atre & Singh, 2016). The way to categorize data on the basis of initial rules, event logs and fuzzy logic constitutes security policy.

5.2 Why Process Mining for Intrusion Detection?

Using anomaly, signature or hybrid based intrusion detection techniques, the big amount of interesting models proposed in the literature to detect intrusion, are built with Artificial Intelligence, particularly thanks to classic Machine Learning and data mining models: Neural Networks, SVM, Decision tree, KNN.

Process Mining can be more useful for Host Intrusion Detection and thus, can easily help to eradicate attacks for instance the 16th one listed in table 1, critical data access and modification. The reason is simple: (Proctor, 2000) The interest of Network based Intrusion Detection is on the analyze of packets within the network while Host based Intrusion is orientated on logs, the starting point of Process Mining.

Process mining theory, for intrusion detection is more interesting than previous Artificial Intelligence models because its models of detection is not built with data examples, but with the security policy. It assures that Intrusion that are founded are really events that violate security policy. Then, alarms are only generated for real attacks, it solves the problem of false positive. But, at the same time, the use of Process Mining increases the rate of false negative if the rules are not enough to consider the different cases of security violations. This last issue
can be addressed by a good definition of important rules for the enterprise, because the most important for an enterprise is not to provide the guaranty of 100% of security, but to implement security mechanisms accordingly with the strategy of the enterprise.

6. Conclusion and Future Works

Information systems are usually the cible of various attacks, that can be made by internal or external attackers. Intrusion detection within such systems is one of the major issues tackled to manage security. Thus, many researchers continue to provide various solutions, for the improvement of existing ones. This paper shows a panoramic view of intrusion detection domain starting by presenting techniques used and most popular attacks. There still exists several relevant problems in intrusion detection depending on the type of considered IDS. This review presents Process Mining as a solution to tackle Host Intrusion Detection challenges, because it improves classic Machine Learning and Data mining technics used before by building the model of detection not on the basis of data examples, but on the basis of rules generated by the enterprise accordingly to its strategy.

References


Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).