ABSTRACT: Today the security aspect is a backbone of every computer system, particularly when this system contains sensitive information. When we use our computers and connect to the network or use web services, we trust that our personal and sensitive information (images, passwords, credit card number …etc) are confidential and well secured by the used security system. Unfortunately, the hackers may succeed to gain access to the system and misuse our information. Generally, the aim of the intrusion is to find new ways to annoy, steal and harm parts of computer systems. Moreover, with the possibility of connecting several computers and networks, the necessity of protecting the whole data and machines from attackers (hackers) that try to get some confident information to use for their own benefit or just destroy or modify valuable information was born. At this point IDS appears to help users, companies or institutions to detect when they are getting compromised. In this paper, we develop an intrusion detection system using multi agent system, ontological techniques (IDSMAO) and misuse method.

Subject Categories and Descriptors: [D.4.6 Security and Protection] [COMPUTER-COMMUNICATION NETWORKS]: Security and protection [I.2.11 Distributed Artificial Intelligence]: Multiagent systems

General Terms: Intrusion Detection Systems, Multi-agent Systems, Semantic Web

Keywords: Intrusion Detection Systems, Computer and Network Security, Semantic Web, Ontology, Multi-agents System (MAS), JADE and JENA

Received: 21 October 2018, Revised 10 December 2018, Accepted 24 January 2019

Review Metrics: Review Scale- 0/6, Review Score- 4.22, Interrevievwer Consistency- 82.5%

DOI: 10.6025/jdim/2019/17/3/133-144

1. Introduction

With the evolution of Internet and computer networks, security has become a major concern over the years. There are many systems of protection against security concerns such as firewalls that protect a network by preventing intrusions from the Internet and control the flow of data going in the network. But these systems cannot prevent all the malicious traffic and they may allow the passing of an intruder. Thus, IDS (intrusion detection systems) have to provide a way to catch all kinds of malicious actions. With the increasing number of attacks on computer systems through web services and applications day by day, there is a high need to protection of our computers. These attacks that occur are basically those that take an advantage of the vulnerabilities and misconfigurations on the systems.

Even with the presence of firewalls and proxies, the attacker is successful in detecting the vulnerabilities and gaining access into the system. The objective of Intrusion Detection Systems (IDS) is to detect and initiate a quick response when an unauthorized activity takes place. IDS prove to be an integral part of every computer system. Typically, when they are used with: security policy, vulnerability assessments, data encryption, user authentication, access control, and firewalls, they can greatly enhance network safety.

In this paper, we propose a new way to detect intruders. We introduce a new IDS, called IDSMAOnto (Intrusion Detection System based on Multi-Agent and Ontology). IDSMAOnto is based on the integration of the multi-agents
system technology and the ontology techniques. The basic functionality of our intrusion detection system is the security guard, it generates an alarm when it detects something suspicious and then the administrator of the network security investigates the cause of the alarm.

Our proposed system uses a set of agents that can be applied to a number of tasks, namely: data capturing, filtering the captured data, also analyze data to detect attacks and ultimately alerting the administrator (as in [4][12]). IDSMAOnto efficiently uses misuse detection strategies as in [14]. IDSMAOnto creates new ontology of the signature of the susceptible attack (collected data), then it compares this signature with the ontology of attacks signatures (mapping process) to decide whether it is an attack or not. Moreover, the power and utility of the ontology is realized by the fact that we can express the relationships between collected data and use those relationships to deduce that the particular data represents an attack of a particular type. Ontologies provide powerful constructs that include machine interpretable definitions of the concepts within a domain and the relations between them [18][9].

The remainder of this paper is organized as follows: in the next section, we present the related works then we introduce the intrusion detection field by giving definitions and background on security approaches and types of IDS. Moreover, we express the advantages and the importance of using multi-agent systems in the field of IDS. In section four, we present our general modeling of the proposed approach. Section five, is dedicated to explain the detailed IDSMAOnto system architecture and our ontology. In the section six, we present the functioning of our system using AUML diagrams. The validation and evaluation of and the results of the developed system are shown in the next section. We finish this paper by a conclusion.

2. Related Works

Security is a focal aspect of every computer system and so the quality of these systems depends on the provided functionalities as well as the degree of security. Several works presented many approaches and methods to ensure high degree of security.

The authors of [13] proposed the ontological model for representing intrusion detection and prevention events, over multi-agents architectures and using intelligence computing for reasoning, classification and pattern recognition. The authors focused first the Intrusion Detection Messages Exchange Format (IDMEF). The aim is to enable interoperability between heterogeneous IDS (IDMEF based on XML). To provide reasoning capability, the authors added Semantic Web Rule Language (SWRL) to allow writing rules expressed in OWL concepts providing reasoning capabilities. The attacks were created with testing tools as Metasploit, IDSWakeUP, Mucus synchronized with Snort. The raw data was converted to XML then processed to OWL instances; furthermore the ontology is updated via SPARQL. The use of ontology is to represents the signatures for known attacks and novel attacks, and the intelligent behavior uses the inference model and reasoning tools integrating neuronal networks in the multi-agent system. According to the authors, the detection accuracy of the proposed ontology based IDS is superior to that obtained for traditional signature based IDS, because it provides better performance in processes such as knowledge representation, cooperation, distribution, intelligence reasoning, reactivity, among others.

The authors of [5] used intelligent combination of clustering technique and an ontology to improve new multi-agents IDS. The proposed system is a new distributed intrusion detection system called OCMAS-IDS (Ontology and Clustering based Multi-AgentS Intrusion Detection System). This system takes the advantages of the multi-agents technology and the benefits of semantic relations as well as the high accuracy of the data mining technique. In this paper, the proposed system uses a set of agents that can be applied to a number of tasks, namely: data capturing process, detecting the known and unknown attack categories and ultimately alerting the administrator if any attack is detected. OCMAS-IDS expressed the advantages of integration of the multi-agents technology, the ontology and the clustering technique. The experiments of this system on a real-life network traffic and a set of simulated attacks, showed the effectiveness of the proposal in terms of (i) the scalability and (ii) the detection ability of our system.

In the paper of [21], design and implementation of ontology based knowledge representation for a peer-to-peer MultiAgent Distributed Intrusion Detection System (OntologyBased MADIDS) are introduced. The aims is to develop a distributed intrusion detection system, by taking advantage of ontology technique to overcome some knowledge sharing limitations of current IDSs. The primary focus of this paper is on the detection of attack tools by implementing Outbound Intrusion Detection (OID), but the aims of Ontology-Based MADIDS not only detect intrusion on a single host, but also in a distributed domain, also with the help of P2P architecture. So if one host detects and adds a new attack to the ontology, it shares it to the other hosts. This implementation makes the framework more flexible and robust and to enable more intelligent behavior in agents. The authors cover different types of Network Attack, which are DoS, R2L, U2R and Probe, also this system is tested on the attack R2L, but never implemented in real cases.

3. Theoretical Foundations

3.1. Computer and Network Security

Nowadays computers and the Internet are used almost in every part of our lives: saving all the data, purchases, office, electronic transactions, learning …etc. With the possibility of connecting several computers and networks was born the necessity of protecting all this data and machines from attackers (hackers) [8]. Computer secu-
curity is the process of preventing and detecting unauthorized use of your computer. Prevention measures help to stop unauthorized users from accessing any part of our computer system. In addition, detection helps to determine whether someone attempted to break into our system, if they were successful or not, and what they may have done [1]. Computer security infrastructure is based on the following three main security services: confidentiality, integrity, and availability in a computer system. Confidentiality is the intruder has access to confidential information and Integrity means that information can be modified or altered by the attacker. Availability means that the system gets blocked so it cannot be used normally [8].

3.2. Vulnerabilities
Vulnerabilities in network security can be summed up as the "soft spots" that are present in every network. The vulnerabilities are present in the network and individual devices that make up the network. Networks are typically plagued by one or all of three primary vulnerabilities: technology weaknesses, configuration weaknesses and security policy weaknesses: [2][11] (they are primary but are not just limited to these three factors).

- **Technological Weaknesses**: Computer and network technologies have intrinsic security weaknesses. These include TCP/IP protocol weaknesses, operating system weaknesses, and network equipment weaknesses.

- **Configuration Weaknesses**: Network administrators or network engineers need to learn what the configuration weaknesses are and correctly configure their computing and network devices to compensate.

- **Security Policy Weaknesses**: Security policy weaknesses can create unforeseen security threats. The network can pose security risks to the network if users do not follow the security policy.

3.3. Intrusion Detection Systems
Several terms are associated in the security that is related to Intrusion are as follows:

- **Intrusion**: Any set of actions that attempt to compromise the integrity, confidentiality, or availability of a computer resource [3]. An intrusion consists of a number of related steps performed by the intruder that violate a given security policy [19].

- **Intrusion Detection**: Intrusion detection is a set of techniques and methods that are used to detect suspicious activity (actions that attempt to compromise the integrity, confidentiality, or availability of a computer resource) both at the network and host level [10]. More formal definition: it relates to the problem of identifying individuals who are using, or attempting to use a computer system without authorization (crackers) and those who have legitimate access to the system but are abusing their privileges (the insider threat) [20].

- **Intrusion Detection Systems**: Intrusion detection systems help computer systems to prepare and deal with attacks. They collect information from a variety of vantage points within computer systems and networks, and analyze these information for symptoms of security problems [19][6][3].

3.4. Analysis types
Intrusion detection systems must be able to distinguish between normal and abnormal activities in order to discover malicious attempts in time:

- **Misuse Detection**: The Signature Based Detection compares a possible threat with the attack type already stored in the IDS. The limitation of this type of detection technique is that if any new type of threat comes which is not already known to the IDS, the system becomes vulnerable to that attack [23]. Misuse detectors are very effective at detecting attacks without generating an overwhelming number of false alarms.

- **Anomaly Detection**: In this type of detection, the IDS looks for vulnerabilities based on rules set forth by the user and not on the basis of signatures already stored in the IDS. This type of detection usually uses Artificial Intelligence to distinguish between normal traffic and anomalous traffic [23]. IDSs based on anomaly detection detect unusual behavior and thus have the ability to detect symptoms of attacks without specific knowledge of details. However, anomaly detection approaches usually produce a large number of false alarms due to the unpredictable behaviors of users and networks [24].

4.5. Multi-agent systems (MAS) and Ontology based IDS
Multi-agent systems (MASs) are systems consisting of more than one autonomous agent that are able to interact with one another. The particular characteristic here is that in order to achieve their goal(s), these agents must coordinate their actions [7]. The multi-agent system has become an increasingly powerful tool in developing complex systems that take advantages of agent properties. Originally, multi-agent systems came from the field of artificial intelligence (AI). At first, this field was called distributed artificial intelligence (DAI) [22][17]. Moreover, Multi-Agent Systems (MAS) is a field of very active research. It has become a new paradigm provides a very suitable architecture for a design and implementation of dynamic open system such as: Security, e-learning, Ecommerce …etc. With agent-based technologies, a support for complex information systems development is introduced by natural decomposition, abstraction and flexibility of management for organizational structure changes. A MAS is a community of autonomous agents working together, sometimes complex ways of cooperation, coordination, competition, in order to achieve a goal: solving a complex problem. The MAS paradigm is a very appropriate platform for IDS and extracting relations and meaning of concepts using ontologies.
Ontology-based IDS solutions are becoming increasingly used in information security. Raskin in [15] developed an ontology for the data integrity of web resources and advocated the use of ontologies for information security. Ontology is the explicit specification of the conceptualization of a domain which captures its context. Ontological models are flexible in defining, nearly any concept to the desired level of detail. They are also fairly easy to extend and the logical foundations of modern ontology languages allows reasoning over concepts within a knowledge base. Reasoning over the instances of the data within the domain can be used to infer new knowledge and to integrate. data. Moreover, ontological models can be easily shared, refined and reused between entities in a domain [9].

4. Modeling of the Proposed Approach

Currently, there are various kinds of intrusions and attacks that can damage computer security. Therefore, we focused on intrusion detection system because it is becoming in important package in security system. Our benchmark is constituted using multi-agent system and formal ontology in order to facilitate the communication between agents.

Figure 1. The proposed model of our approach

Figure 2. Detailed architecture of IDSMAO
IDS is composed of different cooperative, communicant and collaborative agents for collecting and analyzing massive amounts of data represented in formal ontology. However, the ontology is a large among of data in order to define intrusion. Thus, we propose our distributed architecture that contains various agents for collecting and analyzing the data in the environment. These agents work in different levels: capture level, pre-treatment level, analysis level, and reaction level. The agents collect and analyze the data from the environment to define the attacks exploited by the intruders. Our system contains the following agents:

• **The Sniffer Agent**: It is the primer agent, that captures (catches) packets and events from the network, and sends them to other agents to be Filtered, and processed.

• **Filter Agent**: It receives data from Sniffer Agent then it collects packets and filters them.

• **The Misuse Detection Agent**: It analyzes the collected and filtered packets, and compares the found signatures with the predefined patterns.

• **Reaction Agent**: It manages the events to generate reports and logs to inform the user about the intruder.

Finally, the Misuse Detection Agent sends message to sniffer agent to Process the next packet. IDSMOnto detects the known attacks through the intelligent agent MisuseAgent which uses ontology to enrich data intrusions and attack signatures by semantic relationships.

5. Detailed Architecture of our Proposed System

This section describes the proposed model by making use of ontology exploited by a set of intelligent agents.

5.1. Sniffer Agent (Sniffer_A)
Sniffer Agent or capture agent records all the traffic over the network and displays them, like network connection and records all the system events such as system file operation. The sniffer agent has two-class System_Sniffer and Network_Sniffer. First, the sniffer network is responsible for recording all traffic that pass over the network and stores them in the memory. It records several attributes about network: IP address, source, and IP address of destination, whether the host has a successful connection, length of packet (the number of seconds), type of the protocol (TCP, UDP, ICMP, etc.) and the port number …etc. Second, system sniffer is responsible to record the events of the host system (operating system, applications …etc).

Moreover, Sniffer_A can be distributed over the host (network and system) and duplicated to get more information. Sniffer agent gets all the list of interfaces, after that the user chooses the needed interface; then the sniffer agent opens the device (interface) and starts to sniff and get the packets from the network and sends it to Filter_A.

In addition, the important work of this agent is stores these information of the network and the system in a Record File.

5.2. Filter Agent (Filter_A)
The filter agent receives the data from sniffer agent or it can read it from Record File. The main role of Filter Agent is distinguishing the various types of the collected events and traffic: destination address, used protocol, packet category (TCP, UDP, ICMP, etc.); because each kind of information is concerned by a specific kind of intrusion. This task is important because it helps Misuse_detection_A to define the intrusion and facilitate the detection in time. The Filter_A performs its tasks as a pre-treatment phase, which precedes the analysis phase carried out by the following agent.

5.3. Misuse Detection Agent (Misuse_detection_A)
Misuse Detection Agent or analyzer agent, analyzes the packets captured by the Sniffer_A and then pre-processed by the Filter_A. In fact, Misuse_detection_A searches for attack signatures in these packets. When Misuse_detection_A detects an intrusion or attack, it sends a notification to Reaction Agent. This agent cannot detect the attacks without their Signatures, so the signature Ontology was defined. This ontology contains many attack signatures, which are used to discover intrusions or attacks through (mapping the signatures of filtered packets with the signatures ontology).
In the case of 100 % similarity between the events ontology and any signature attack ontology, the system alerts the user and gives the necessary information about the attack (type, addresses, port ...etc). But, if there is not total similarity (100 %), we proposed the shortest path method (simple and efficient method in our case) to calculate semantic similarity based on the ontology hierarchy, suggesting that the shortest path between two nodes was the simplest approach for measuring distance between two terms [16]. In mathematics, the formula for the distance between two nodes by the shortest path was denoted by,

\[\text{Sim}(c_1, c_2) = 2 \times \frac{\text{MAX} - L}{\text{MAX}}\]

where \(c_1\) and \(c_2\) were the compared nodes, \(\text{MAX}\) the maximum path on the hierarchy, and \(L\) the shortest path. The main advantage of this method was its low complexity in calculation. Moreover, when only the is-a relationship existed in a semantic network, semantic relatedness and semantic distance were equivalent. However, this method was short of consideration for different kinds of edges as well as the semantic relatedness representing these edges. According to the Similarity degree, the misuse agent alerts the user by the susceptible attack and gives the degree of similarities with the different attacks.

5.4. Reporter (reaction) Agent
To implement passive response to security incidents, the Misuse Detection Agent report its results to the Reaction Agent. Whenever an intrusion is detected, Reaction Agent sends an alert to the system (or user). This alert can be a message popup on the screen or a message to a centralized machine or an alert file. The result record contains the information including source IP (Source Host), type of the suspicious attack (Attack Type) ...etc. In addition, the Report_A decides the response type to implement it according to the detected event.

5.5. Ontology
Our ontology represents our database system. We adopt ontology technique to represent intrusion and detection knowledge (attacks signatures). Ontology has been developed using the detailed description of all-important security concepts and attacks signatures. Several attacks are represented by the ontology and we use this ontology to provide detection process (mapping process) to inform the user to take the appropriate actions. The power and utility of the ontology is realized by the fact that we can express the relationships between collected data and use those relationships to deduce that the particular data represents an attack of a particular type.

After collecting the information about the received packet, Misuse detection agent forms the ontology of the packet (representing this packet). Then, this agent performs the process of mapping (calculate the similarity) between the created ontology (packet ontology) and the ontologies that represent the attacks signatures as it is explained in the section 5.3.
The PacketCollection class shown in the figure 7 is used to group common packet instances and classify them according to the type of the packet (PacketType class). For instance, if there were multiple ping packets to the same node within a specific time frame, an instance was created in the PacketCollection class of PacketType and PingFloodType. These instances were later used by IDSMAOnto to assist the attack identification.

5.5.2. Representation (Attacks signatures ontology)
The attack ontology (Figure 8) is used to provide information about simple and complex attacks. The attack data is obtained by using inference through ontology constructs and rules. The figure 8 (attack ontology) presents a part view of our ontology. The Class thing (representing the class of all things) contains class Attackpacket has subclass Attack. This latter class is comprised of the following subclasses:

- **Recon:** Refers to Reconnaissance, it means gathering information. This attack has intention to gather information by using scan technique such as: scan TCP or sniffing …etc.

- **GainAccess:** Not having official permission or approval to the network. This includes the following subclasses: privilege gains or unauthorised access.

- **ViewChangeData:** This class contains attributes representing malicious code that aim to change data (alteration).

- **Availability:** This class contains attributes representing attack DOS and spoofing.

The figure 9 shows the Availability classes. There are two primary techniques an attacker may use to make a host or network unavailable. These two techniques are a denial of service or spoofing attack. An attacker gains unauthorized access to a computer or a network by making it appears that a malicious message has come from a trusted machine. There are two primary techniques using spoofing ARPspoofing and IPspoofing.

The class DOS (figure 10) contains attributes representing particular attack DOS (Denial of Service). DOS attack has two subclasses CrashNode and Resources. These classes represent more than one type of attacks. For example, a SYN flood, smurf, teardrop, ping–of–death.

Figure 11 illustrates ping of death attack using owl language, it has packetLenof 65535 and protocol is ICMP.

Our system uses another type of attack (figure 13): complex attack. Complex DOS has a set of attributes that represent simple attacks: a Ping scan, Node scan, and Availability attack.

The system is designed to use attack ontology and the ontology of captured traffic (calculation of similarity) to define the susceptible attacks. This allows better adaptability and flexibility in attack detection. The majority of the customized code was to initially populate the knowledge base with traffic data using a mapper program developed using Java and Jena.
6. Functioning of our System: AUML Diagrams

In this section we present AUML diagrams representing the functioning and the structure of our system.

6.1. Sequence Diagram
The sequence diagrams in AUML represent message exchanges between agents. The sequence diagram given in the figure 13 shows the interaction between different agents:

- The sniffer agent collects all the events occurred in the host to which it is related and stores the collected data in a sniffing file and sends it to Filter Agent.
- The Filter agent will treat these crude events and send it to missus detection agent for the analysis process.
- The missus detection agent receives data from Filter Agent, and compares signatures with predefined patterns in the attack ontology using mapping process. If there is intrusion, this agent sends the necessary data to Reaction Agent.
- Reaction Agent tries to fix the problem or sends an alert to the user.
6.2. Activity Diagram of Data Capturing
Figure 14 shows the data capturing activity. The sniffer agent gets the list of all interfaces and its corresponding network devices and displays, to the user, the interface name and obtains the interface descriptions. Then, it gets network interface IP and subnet mask.

The user chooses the network interface for capturing. The agent starts listening to network and gets every packet that crosses the network. After capturing the packets, the agent sends them to be filtered and analyzed by filter agent and analyzer agent. The system saves all the packets information in the signatures file.

6.3. Analysis Activity Diagram
Figure 15 shows the analysis activity diagram. The system saves the captured filtered data in the existing signatures file (or create new file). The analyzer agent checks and defines abnormal system behaviors and normal behaviors using the mapping process of the stored signatures and those predefined signatures (attack ontology).
7. Implementation and Results

For the validation, we have used several programming languages and environments. Eclipse is an integrated development environment used in computer programming. WinPcap is an open source library for packet capturing and network analysis. Most networking applications access the network through widely used operating system primitives such as sockets. It is easy to access data on the network with this approach since the operating system copes with the low-level details (protocol handling, packet reassembly, etc.). Jpcap is an open source library for capturing and sending network packets from Java applications. It provides facilities to: Capture raw packets live from the wire, save captured packets to an offline file, and read captured packets from an offline file, filter the packets according to user-specified rules before dispatching them to the application, send raw packets to the network. A Java Development Kit (JDK) is a development environment for writing Java applets and applications. It consists of a runtime environment that "sits on top" of the operating layer as well as the tools and programming that developers need to compile, debug, and run applets and applications written in the Java language. In addition, we used Visual Layout Tool for JavaFX Applications.

JADE (Java Agent Development Framework) is a software Framework fully implemented in the Java language. It simplifies the implementation of multi-agent systems through a middle-ware that complies with the FIPA specifications and through a set of graphical tools. Apache JENA is an open source Semantic Web framework for Java. It provides an API to extract data from and write to RDF graphs. The graphs are represented as an abstract "model". A model can be sourced with data from files, databases, URLs or a combination of these. A Model can also be queried through SPARQL. FXML is a scriptable, XML-based mark-up language for constructing Java object graphs. It provides a convenient alternative to constructing such graphs in procedural code, and is ideally suited to defining the user interface of a JavaFX application, since the hierarchical structure of an XML document closely parallels the structure of the JavaFX scene graph. RDF is a directed, labeled graph data format for representing information in the Web. RDF is often used to represent, among other things, personal information, social networks, metadata...etc. SPARQL is an RDF query language, that is, a query language for databases, able to retrieve and manipulate data stored in Resource Description Framework format. Our main window is shown below:

![Figure 16. Principal interface](image)

This interface allows the user to show statistical information about the type of address consuming IPv6/IPv4/ARP. There are counters for each of the three protocols: ICMP, UDP and TCP. In addition the user is given the ability to count the traffic for any number protocol. Also, the amount of data received and sent is shown per protocol.

In the next figures, we present some query results. These results provide the information requested by SPARQL. The ontology API Jena uses an ontology model. The specific ontology model used by IDSMAOnto is the OWL_MEM_RULE_INF model. This specific ontology model supports the OWL full language. Our IDSMAOntology is implemented in real time detection. In the figure 18, we present the physical addresses that the host is connected to them.
Figure 18. Sample query results from the IDSMAOnto

Figure 19 presents a portion of the captured packet during the intrusion detection process.

Using several tests (Ping of death, service denial, flooding, spoofing …etc), our application detected all of them and showed many alerts (most of them are correct alerts; rarely, we receive incorrect alerts because of the quality of the used equipment and software). But we aim to use it in a real case (real attacks in a network).

8. Conclusion

Using ontology with Multi-Agent System and reasoning models in the Intrusion detection systems (IDS) is a promising and efficient approach. The detection architecture based on reaction rules generated by the intelligent and correlated component in our IDSMAOnto creates new reaction rules in other security components. Ontology of attacks (signatures) and communication protocols provide a powerful construct for improving the detection capability of the application. Various instances of attacks and corresponding vulnerabilities are tested. This work presented an ontological model integrated with multiagent system using artificial intelligence that provide better performance in processes such as knowledge representation, cooperation, intelligence reasoning, reactivity…etc. The use of IDSMAOnto is a valuable asset to a network manager, it allows him to understand the weaknesses in the network and take corrective action to prevent future attacks.

In future work, we aim to work on the detection of other types of attacks such as: DDoS and variants of the DoS, SQL Injection and others, and extend the number of tests and real case study.

References


