

# Ontology-based Distributed Intrusion Detection System

F. Abdoli\* and M. Kahani\*\*

Ferdowsi University of Mashhad, Mashhad, Iran  
Communication and Computer Research Center, Mashhad, Iran

\* Email: fateme.abdoli@gmail.com

\*\*Email: kahani@ferdowsi.um.ac.ir

**Abstract**—In this paper we discussed about utilizing methods and techniques of semantic web in the Intrusion Detection Systems. To extract semantic relations between computer attacks and intrusions in a Distributed Intrusion Detection System, we use ontology. Protégé software is our selected software for building ontology. In addition, we utilized Jena framework to make interaction between MasterAgent and attacks ontology. Our Distributed Intrusion Detection System is a network which contains some IDSagents and a special MasterAgent. MasterAgent contains our proposed attacks ontology. Every time a IDSagent detects an attack or new suspected condition, it sends detection's report for MasterAgent. Therefore, it can extract the semantic relationship among computer attacks and suspected situations in the network with proposed ontology. Finally, the experience shows that the proposed system reduced the rate of false positive and false negative.

**Keywords** — Ontology; Intrusion Detection System; Denial of Service attack

## I. INTRODUCTION

Since intrusion detection was introduced in the mid-1980s, Intrusion Detection System (IDS) has developed for almost twenty years to enhance computer security. High false negative and false positive prevent using intrusion detection system practically. In these years the computer scientists attempt to solve this problem and reduce these false rates. They use so many methods and techniques to improve these IDS systems, such as: Data mining [13], State Transition diagrams [12], Clustering [11], Classification [10] and Neuro-Fuzzy methods [14] .... And try to reduce false rate and increase their reliability.

Semantic web techniques and methods like concept of “content” and “ontology” can be used in many fields of computer science. Victor Raskin et al. [4] opened new field of Information Security, they discussed about using “Ontology” in Information Security and its advantages. They believed that ontology is an extremely promising new paradigm in Computer security field. They say by using ontology we have a strong classification tools for unlimited events.

Every information security method which can use the concept of “content”, it can utilize methods and techniques of semantic web. Intrusion Detection Systems

is a good example. In this field some researches has done recently. Those researches try to improve Intrusion Detection System by utilizing ontology in their systems. The main goal of this paper is to use these methods in order to improve Intrusion Detection System and reduce their false rate.

The remainder of the paper is to organize as follows: Section 2 presents related work in the domain of using semantic web methodology in the Intrusion Detection Systems. Section 3 presents our proposed model and proposed ontology and at the end of this section we will discuss about the advantages and disadvantages of our proposed system. Section 4 is about our future work, and finally conclusion is in the end section.

## II. RELATED WORK

Using semantic web methodology in Intrusion Detection Systems is new. The first research in this area was done in 2003 [1], [2]. So far there is little research that was done in this domain and they utilize different usage of ontology in this area.

The first research was done by Jeffrey Undercoffer and et al. [1], [2]. They produced an ontology specifying a model of computer attack. Their ontology is based upon an analysis of over 4,000 classes of computer intrusions and their corresponding attack strategies and it is categorized according to system component targeted, means of attack, and consequence of attack and location of attacker. They argue that any taxonomic characteristics used to define a computer attack be limited in scope to those features that are observable and measurable at the target of the attack. They present their model as a target-centric ontology. Moreover, specifying an ontological representation decouples the data model defining an intrusion from the logic of the intrusion detection system. The decoupling of the data model from the Intrusion Detection System logic enables heterogeneous Intrusion Detection System's to share data without a prior agreement as to the semantics of the data.

The second work presents ontology to describe relationship among features observed by multi-senor. There exist two kinds of nodes in ontology value nodes and attribute nodes. By assigning the weight to the edge

between values nodes and their parent attributed node, they provide more flexible matchmaking method for intrusion detection. At the same time, the relationship between attribute nodes and their parent can indicate the locality of desired information. An ontology based cooperative detection function is also given in this work [6].

The third work has described an ontology-supported Outbound Intrusion Detection architecture that organizes agents into execution sub-environments called agent cells. The peer-to-peer arrangement of the cells provides a robust non-hierarchical agent structure, and the cells themselves constitute a way of dealing with the malicious-host problem. An attacker-centric ontology serves as a common-knowledge layer for all agents. Traffic and process signatures are generated and matched by independent cells that provide full intrusion detection functionality. Corrugators fuse diagnosis from multiple cells in order to provide more accurate detection. Here the Intrusion Detection architecture ontology is necessary to enable more intelligent behavior in agents, to optimize communication contents and interpretation. Also to give formalism to the way the components of an architecture interact [3], [7].

In another research, the researcher use semantical ontology for security domain of Intrusion Detection System. And by utilizing that ontology, raw alarms come from heterogeneous Intrusion Detection System, integrated. In fact this method can be used for extracting attack scenarios [5].

In the last project the authors propose a novel Breadth and Depth Bayesian classifier and an inference probabilistic algorithm. The inference algorithm is applied over well defined conceptual information integrated in a hybrid Intrusion Detection System by means of ontologies. They said that trying to combine both semantic modeling and probabilistic modeling might be exploited for attack prediction in the Pervasive Computing paradigm [9].

In many field of the Intrusion Detection System we can use the semantic web techniques, for example, we can use them in the domain of analyzing user behavior and system activities or identifying known attack pattern, and also field of analysis of abnormal behavior and activity of systems and etc. In this paper we utilize the concept of ontology to extract semantical relationship among attacks, intrusions, and suspect activities that occur in different systems in our network and attempt to reduce false rate in Intrusion Detection Systems.

### III. PROPOSED MODEL

#### A. Introduction of the proposed model

The most deficiencies of current Intrusion Detection Systems are high False Negative and False Positive. Occasionally Intrusion Detection Systems have mistake in their detections and sometimes the inefficient detection is partly caused by insufficient audit data. Some of Intrusion Detection Systems depend on only one kind of sources network data or host data. However many intrusions can show characters in both of two data sources. For these

reasons we can extend analyzing and investigating field and utilizing more than one system for detecting intrusions and suspicious activities in the network.

Also using of new methods and techniques in the detection phase can reduce the false rate. The use of "Ontology" and the utilization of computer attacks ontology which demonstrated semantic relations between attacks and intrusions is a new solution for attacks detection. It can be use in every IDS. In this paper we utilize this method. The ontology [23] can be seen as an abstraction of a computer-based lexicon, thesaurus, glossary or some type of structured vocabulary, suitably extended with knowledge about a given domain. The domain ontology can be considered as a representation of a domain conceptualization describing possible concepts and relationships between these concepts.

The aim of this research is to present Ontology based Distributed Intrusion Detection System (ODIDS), therefore we implemented multi agents Distributed Intrusion Detection System. We have two kinds of agents in the proposed system, some IDSagents and one MasterAgent. Every IDSagent acts the same as an Intrusion Detection System and they can report their suspicious and malicious status for the MasterAgent. MasterAgent is the most important agent in the multi agents system, its equipped with the proposed attacks ontology. Whenever one of the IDSagents detects a suspicious status, it provides a report and then sends this report to the MasterAgent.

It is noted, that each of the IDSagent can be designed to act like special network-based or host-based Intrusion Detection System. Because of one of the primal goals that we had, was to extend analyzing field for improving the final intrusion detection accuracy and reliability. But for the simplicity we use one kind of IDSagent. Designed IDSagents act like a network-based Intrusion Detection System. Their task is like network intrusion detector, and their database of reports contains a standard set of data to be audited, which includes a wide variety of intrusions. Whenever each IDSagent wants to send a report to the MasterAgent, it randomly selects one of its stored reports and sends it for the MasterAgent. It means that they detect system's situation and do the suitable reaction and in addition they send proper report for the MasterAgent.

Whenever MasterAgent receives a report from other IDSagents, it peruses the detected situation and found its status in existing attack's ontology. In the next part of this section we will discuss about the proposed ontology.

For example, the MasterAgent receives a report that contains some information about the connections in the network of IDSagent. The information is about these topics:

- Network service on the destination
- Type of the protocol
- Normal or error status of the connection
- ...

The MasterAgent peruses the received report, and in order to extract information, it queries the Attacks ontology. After analyzing the detected situation in current

ontology, the MasterAgent can find out one of these results:

- Received report shows DoS attack.
- Received report doesn't show DoS attack.

It means that MasterAgent categorizes the network connections to DoS and notDoS, because our ontology is in the domain of Denial of Service attack.

By querying the Attacks ontology, if received report shows DoS attack, the MasterAgent will find out the type of attack from the attacks ontology.

After studying the status of the detected situation in the Attacks ontology and with regard to the obtained result, the MasterAgent sends suitable alarm to the related IDSagent so that it accomplishes suitable reaction and if it is necessary the MasterAgent can update its ontology.

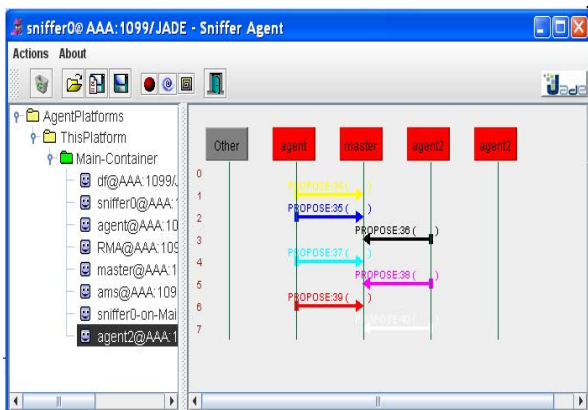


Figure 1. Agents' communication in Jade framework

To implement the multi agent system we use Jade, Java Agent Development Framework. To make relations between MasterAgent and attacks ontology we use Jena, Java framework for building Semantic Web applications, and finally to query the ontology we use SPAQL, query language for RDF and the Semantic Web. Figure 1 illustrates the agent's communications in Jade framework.

### B. Proposed Ontology

There are several ways to build ontology for a special domain. For example, we can reuse old ontology, which is available in that domain. We can rebuild and complete them. Another way to build ontology is using available taxonomy in that domain in order to build related ontology based on that taxonomy [18]. Because of the novelty of using the concept of ontology in the domain of Intrusion Detection system there is few ontologies and they are neither comprehensive nor good for our purpose. For this reason we choose the second way to build our proposed ontology. For this intention we use the taxonomy which is introduced by Hansman et al. [17]. They proposed taxonomy consists of four dimensions which provide a holistic taxonomy in order to deal with inherent problems in the computer and network attack field. The first dimension covers the attack vector and the main behavior

of the attack; in this dimension attacks can be categorized in the following groups: Viruses, Worms, Buffer overflow, Denial of service attacks, Network attacks, Password attacks, and Trojans etc. The second dimension allows for classification of the attack targets, it says the target of an attack, for example, hardware or software. Vulnerabilities are classified in the third dimension and payloads in the fourth.

To find attacks scenario, their behavior, and their effect in the target we studied more than thousands records of the network connections status, which lead to Denial of Service attacks. Most of the records are about Smurf and Neptune attacks. By these studies we found the attacks relationship and modified the properties of the classes of ontology.

To design our proposed ontology we use Protégé software which is free and open source [20].



Figure 2. High Level Illustration of the Proposed Ontology

Figure 2 presents a high level graphic illustration of our proposed ontology. Designed ontology has one main class "attack class". This class contains all kinds of computer attacks and has so many subclasses and branches.

We expand our designed ontology only in the domain of Denial of Service attacks to simplify and reduce the complexity and the amount of its operation. Figure 3 illustrate the Denial of Service class of our ontology.

Every time the MasterAgent receives a report from other IDSagents, with regard to the attack or detected situation properties; if it is necessary, it will update its ontology and send alarm to the related system(s). Every time the central system receives a report as an attack from other systems, with consideration to the attack properties, it updates its ontology. Therefore if it receives a same report as a suspected behavior from other systems, it will easily specify its attack, and also the vice versa process can be done.

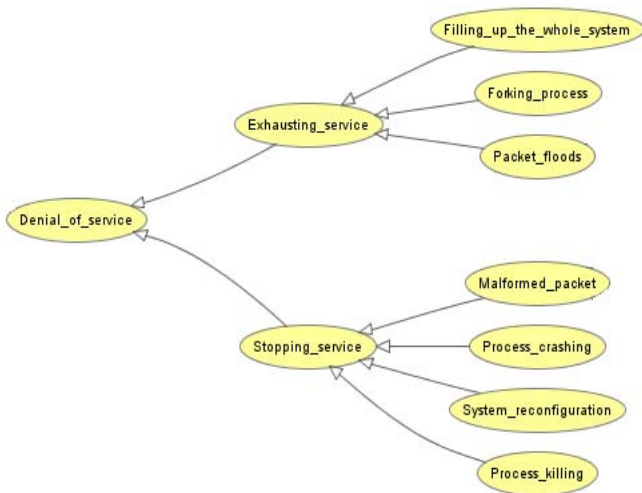


Figure 3. Illustration of the Denial of Service class

### C. Advantage and disadvantage

Our proposed ODIDS can be criticized as bellow:

- single point of failure
- time-wasting

The MasterAgent is a single point of failure. If an intruder can somehow prevent it from working, for example, by crashing or slowing down the host where it runs the whole network will be unprotected. But we assume that in our network every IDSagent has its own Intrusion Detection System. We can solve the single point of failure in our network by this property.

The other criticism of our system is time wasting. Our system needs additional time to make a connection between the MasterAgent and the other IDSagents in the network and to send or receive message among them. We can improve our proposed ontology and also the querying methods of the MasterAgent to facilitate its processing and extracting knowledge to solve mentioned problem.

In our network every agent utilizes two Intrusion Detection Systems. First, its own and the second one is the ontology-based Intrusion Detection System on the MasterAgent. This characteristic is an important property of our network that we hope, it reduces the false rate. In addition, if they found a new attack or suspected situation in their system by using own Intrusion Detection System, the MasterAgent can also help them and send suitable alarm. In the special situation which is unknown and local Intrusion Detection System could not find whether it is an attack or not, maybe the MasterAgent can decide a bout it. Because it has attacks ontology and it can extract the semantic relations among attacks and intrusions. Therefore by studying this ontology the MasterAgent can easily find that whether a special situation is a misuse behavior or not. Thus false negative and false positive are reduced.

## IV. EXPERIMENTAL COMPARISONS AMONG RELATED ALGORITHMS ON KDD 99

Our experimental dataset was the KDD Cup 1999 Data [6], which contained a wide variety of intrusions

simulated in a military network environment. The dataset is about 4 gigabytes of compressed tcpdump data of 7 weeks of network traffic. The simulated attacks fell in one of the following four categories: (1) Denial of Service, (2) R2L — unauthorized access from a remote machine, (3) U2R — unauthorized access to local Root privileges by a local unprivileged user and (4) Probing — surveillance and other probing for vulnerabilities. Our designed ontology is in the domain of Denial of Service attacks. For this reason, we focus on these types of attacks. For example, a SYN flood, smurf, teardrop, ping-of-death, etc. In summary KDD cup99 dataset contains some type of attack and 41 features for each of them.

To identify the performance differences on our ontology-based system and other related algorithms, one measure is the Cost per Example (CPE). It requires two quantities to be defined cost matrix [6] and confusion matrix [6]. A cost matrix (C) is defined by associating classes as labels for the rows and columns of a square matrix in the current paper, there are two classes {DoS, notDoS}, therefore the matrix has dimensions of 2x2. An entry at row  $i$  and column  $j$ ,  $C(i,j)$ , represents the non-negative cost of misclassifying a pattern belonging to class  $i$  into class  $j$ . A confusion matrix (CM) is similarly defined in that row and column labels are class names. An entry at row  $i$  and column  $j$ ,  $CM(i,j)$ , represents the number of misclassified patterns, which originally belong to class  $i$  yet mistakenly identified as a member of class  $j$ . Given the cost matrix with the same cost that is equal to 1 and the confusion matrix obtained subsequent to an empirical testing process, cost per example (CPE) was calculated using the following formula.

$$CPE = \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^m CM(i,j) * C(i,j)$$

Where CM corresponds to confusion matrix, C corresponds to the cost matrix, and N represents the number of patterns tested. A lower value for the cost per example indicates a better classifier model. Table 1 illustrates our CPE results for the ODIDS system.

TABLE I.

CPE RESULTS FOR THE ODIDS SYSTEM

The tested dataset	N	CPE
The labeled 10% KDDcup 99	494.22	0.012
The labeled KDDcup 99	20.00	0.016
The labeled KDDcup 99	740.00	0.017
	0.015	(Average) CPE

The other measure that we use are detection rate and false alarms, which are widely accepted as standard measures

[13,1,2]. With these results we depict ROC<sup>1</sup> diagram for the proposed system. By the ROC diagram we can found the effect of the system's parameters changes on the system's evaluations measures and trace the results for depicting ROC diagram. Figure 4 show the depicted ROC diagram for the ODIDS.

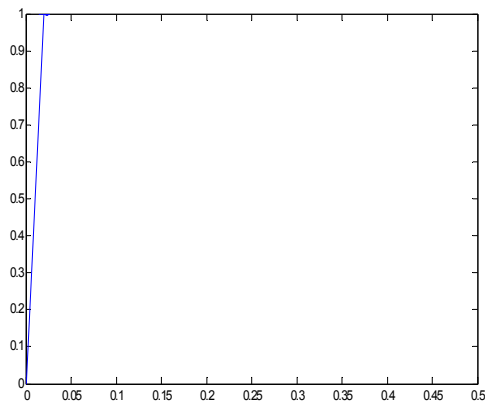


Figure 4. Illustration of ROC diagram for the ODIDS system

In this subsection, we will compare our proposed system with other related algorithms. Experimental results for the ODIDS system and comparison with other algorithm are presented in Table 2. This table shows that ODIDS system demonstrates superior detection performance compared to others. In the case of DoS category, our algorithm detected more than 99.9% of attack records and others did not. It means that ontology-based model is a good candidate for intrusion detection system.

TABLE II.

COMPARISON OF ODIDS SYSTEM AND OTHER ALGORITHM

CPE	FA	DoS	algorithms
0.015	2.5	99.97	ODIDS (our)
0.1579	1.9	99.5	ESC-IDS[22]
n/r	3.5	99.7	RSS-DSS [26]
0.2024	n/r	96.7	Parzen-Window [23]
0.2331	0.6	97.1	Winner of KDD [24]
0.2356	0.6	97.5	Runner Up of KDD [25]
0.2371	0.4	96.9	PNrule [27]

## V. FUTURE WORK

Our future work will focus on the improvement of proposed attack ontology in intrusion detection domain. It means that we want to improve it to contain all kinds of

<sup>1</sup> Receiver Operating Characteristic

attacks. Another future work is to expand the ontology and the ODIDS system to cover distributed attacks. The MasterAgent can extract the semantic relations among attacks and intrusions, therefore in the special situation in one IDSagent maybe it cannot be an attack, but with the other situations that occur on the other agent, they together can make an attack.

## VI. CONCLUSION

This paper introduced a novel Distributed Intrusion Detection System that used special attack ontology to detect attacks and intrusions. For simulating our proposed system we implemented a multi agents system with Jade. This system contains two kinds of agents that the MasterAgent equipped with the designed attacks ontology in this research.

A simulation study was performed to assess the performance of the related algorithms on the KDD 1999 Cup intrusion detection dataset. Simulation results demonstrated that the proposed ODIDS system gets better results. Furthermore, reduction in cost per example was also achieved by using the ontology-based model.

## REFERENCES

- [1] Undercoffer. J, Joshi. A, Pinkston. J, "Modeling Computer Attacks: An Ontology for Intrusion Detection," Springer, pp. 113-135, 2003.
- [2] J. Undercoffer, A. Joshi., T. Finin, and John Pinkston, "A target centric ontology for intrusion detection: using DAML+OIL to classify intrusive behaviors," Knowledge Engineering Review, Cambridge University Press, pp. 23-29, January, 2004.
- [3] S. Mandujano, A. Galván, J. A. Nolzco, "An Ontology-based Multiagent Architecture for Outbound Intrusion Detection", 3rd ACS/IEEE International Conference on Computer Systems and Applications, AICCSA '05, vol. 1, pp. 120-128, Cairo, Egypt, January 2005.
- [4] V. Raskin, C. Helpenmann, K. Triezenberg, and S. Nirenburg, "Ontology in information security: a useful theoretical foundation and methodological tool", New Security Paradigms Workshop, ACM Press, pp. 53-59, Cloudcroft, NM, 2001.
- [5] Yan, W., Hou, E., Ansari, N., "Extracting and querying network attack scenarios knowledge in IDS using PCTCG and alert semantic networks," IEEE International Conference 2005.
- [6] Yanxiang.H, Wei.C, Min.Y and Wenling.P , "Ontology Based Cooperative Intrusion Detection System," Network and Parallel Computing, 2004 springerlink
- [7] Mandujano. S, "An Ontology-supported Intrusion Detection System," Taiwanese Association for Artificial Intelligence, 2005
- [8] Klaus. M, IDS - Intrusion Detection System, 2005
- [9] Anagnostopoulos, T.; Anagnostopoulos, C.; Hadjiefthymiades, S., "Enabling attack behavior prediction in ubiquitous environments," Pervasive Services, ICPS '05. 2005.
- [10] Gomez J., Dasgupta D., "Evolving Fuzzy Classifiers for Intrusion Detection," Proceeding Of 2002 IEEE Workshop on Information Assurance, United States Military Academy, West Point NY, June 2001.
- [11] Guan Y., Ghorbani A. And Belacel N., "Y-means: A Clustering Method for Intrusion Detection," Proceedings of Canadian Conference on Electrical and Computer Engineering. Montreal, Quebec, Canada. May 4-7, 2003.
- [12] Ilgun K., Kemmerer R.A., and Porras P.A., "State Transition Analysis: A Rule-Based Intrusion Detection Approach," IEEE Transaction on Software Engineering, Vol 2, No 3, 21(3), March 1995.
- [13] Lee W., Stolfo S.J., Mok K., "A data mining framework for building intrusion detection models," Proceedings of IEEE Symposium on Security and Privacy, pp 120 -132, 1999.

- [14] Mohajerani M., Morini A., Kianie M. "NFIDS: A Neuro-Fuzzy Intrusion Detection System," IEEE 2003.
- [15] Lait, Leslie R.; Nash, Eric R.; Newman, Paul A. , "The df A proposed data format standard," NASA Center: Goddard Space Flight Center, Mar 1, 1993
- [16] 09-Ashbindu-GEAS\_19 October - The advantage of standard format alerts. [www.oasis-open.org/events/ITU-T-](http://www.oasis-open.org/events/ITU-T-)
- [17] Simon H, Ray , " A taxonomy of network and computer attacks," Elsevier, Computers & Security (2005) 24, 31e43
- [18] Deborah L. McGuinness, Ontology Come og Age, spinning the semantic web, 2003.
- [19] DU.Y, WANG. H, PANG. Y, "Design of A Distributed Intrusion Detection System Based on Independent Agents," IEEE 2004.
- [20] <http://protege.stanford.edu>
- [21] [kdd.ics.uci.edu/databases/kddcup99/kddcup99.html](http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html)
- [22] A. Nadjaran.T, M.Kahani, R.Monsefi , "Network Intrusion Detection Based on Neuro- Fuzzy Classification," ICOCI2006 (Kuala Lumpur, Malaysia, June 6-8, 2006
- [23] Yeung D. Y., Chow C., "Parzen-window Network Intrusion Detectors, Sixteenth International Conference on Pattern Recognition," Quebec City, Canada, pp. 11-15, August 2002.
- [24] Pfahringer B., Winning the KDD99 Classification Cup: Bagged Boosting, SIGKDD explorations, 1(2), 65-66, 2000.
- [25] Levin I., KDD-99 Classifier Learning Contest LLSoft's Results Overview, SIGKDD Explorations, ACM SIGKDD, 1(2) 67-75, 2000.
- [26] Song D., Heywood M.I., Zincir-Heywood A.N., "Training Genetic Programming on Half a Million Patterns: An Example from Anomaly Detection," IEEE Transactions on Evolutionary Computation, 2005.
- [27] Agarwal R., Joshi M. V., "PNrule: A New Framework for Learning Classifier Models in Data Mining," Technical Report TR 00-015, Depratment of Computer Science, University of Minnesota, 2000.