Improve the efficiency of intelligent IDPS by using Reinforcement Learning

Georgi Tsochev, Roumen Trifonov, Radoslav Yoshinov, Slavcho Manolov and Galya Pavlova

***Abstract*—**The present paper describes some of the results obtained in the Faculty of Computer Systems and Technology at Technical University of Sofia in the implementation of project related to the application of intelligent methods for increasing the security in computer networks. Also is made a survey about existing hybrid methods, which are using several artificial intelligent methods for cyber-defense. The paper introduces a model for intrusion detection systems where multi agent systems are the bases and artificial intelligence are applicable by the means simple real-time models constructed in laboratory environment.

***Keywords*—** multi-agent systems, artificial intelligence, network and information security, intrusion detection system, intrusion prevention system, reinforcement learning.

I. Introduction

With the development of cyber-attacks, the human factor is not sufficient for timely analysis and attack action. Human resources and the lack of expertise are the main weaknesses of organizations. The fact is that most network attacks are carried out by intelligent agents, such as computer worms and viruses. Therefore, the fight against them can be done with intelligent semi-autonomous or wholly autonomous agents that can detect, evaluate, and respond with the appropriate protection action. These intelligent methods will need to be able to manage the entire process in response to an attack, to analyze and establish what type of attack is happening, what is the purpose and what is the appropriate response, and last but not least how to prioritize and prevent secondary attacks. It is in these difficult situations that we need innovative approaches such as the application of artificial intelligence methods.

Artificial intelligence (AI) gives us the opportunity to develop more secure applications [14]. Based on that the department of Information technologies in industry at Technical University of Sofia began research of intelligent Intrusion detection/prevention system (IDPS). So far, the system is working with multi-agent technology for enhancing the security of a network, segment of a network or individual host.

Our proposed model consists of two major multi-agent frameworks – host based monitoring system and network gateway monitoring system (partly based on rules) [15]. The two frameworks operate at different layers. The proposed system work is divided into five layers (Fig. 1) – network layer, system hardware, transport layer, data layer and system software. Layers can be merged according to their intended purpose [15]. The first three network, system hardware, transport are grouped together in the TCP / IP stack area, and transport, data, operating system level software (Operating System). As you can see, the transport level is involved in both groups due to the fact that it has an important dual character.

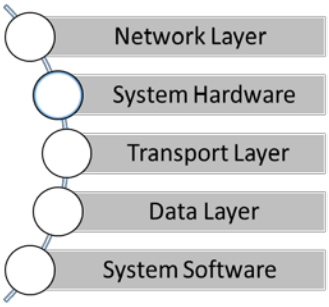

Fig. 1. Operating layers of the Proposed System [15]

As mentioned, many organizations use encryption or similar techniques to conceal the basic information in the body of the package. The proposed system inspected the header to detect traces of penetration of the first two stages of the life cycle of one or more attacks. Some of the protection actions can be one of the following, based on the statistics of the top 10 attacks [16]:

1) Dropping a bundle
2) Blocking
3) End the TCP connection
4) Record.

This research is conducted and funded in relation to the execution of a scientific-research project № H07/56 "Increasing the level of network and information security using intelligent methods" under the contract with National Science Fund in Bulgaria.

Prof. R.Trifonv is a lecturer at the Technical University of Sofia. He is a head of department "Information Technologies in Industry. (e-mail: r_trifonov@tu-sofia.bg).

Assist.Prof. G. Tsochev lecturer at the Technical University of Sofia. He is now with the department "Information Technologies in Industry. (e-mail: gtsochev@tu-sofia.bg).

R. Yoshinov is Director of Laboratory of Telematics at the Bulgarian Academy of Sciences (e-mail: yoshinov@cc.bas.bg).

Assoc. Prof. S. Manolov works at the Technical University of Sofia (e-mail:slav1943@gmail.com).

Assist.Prof. G. Pavlova lecturer at the Technical University of Sofia. She is now with the department "Information Technologies in Industry. (e-mail: raicheva@tu-sofia.bg).

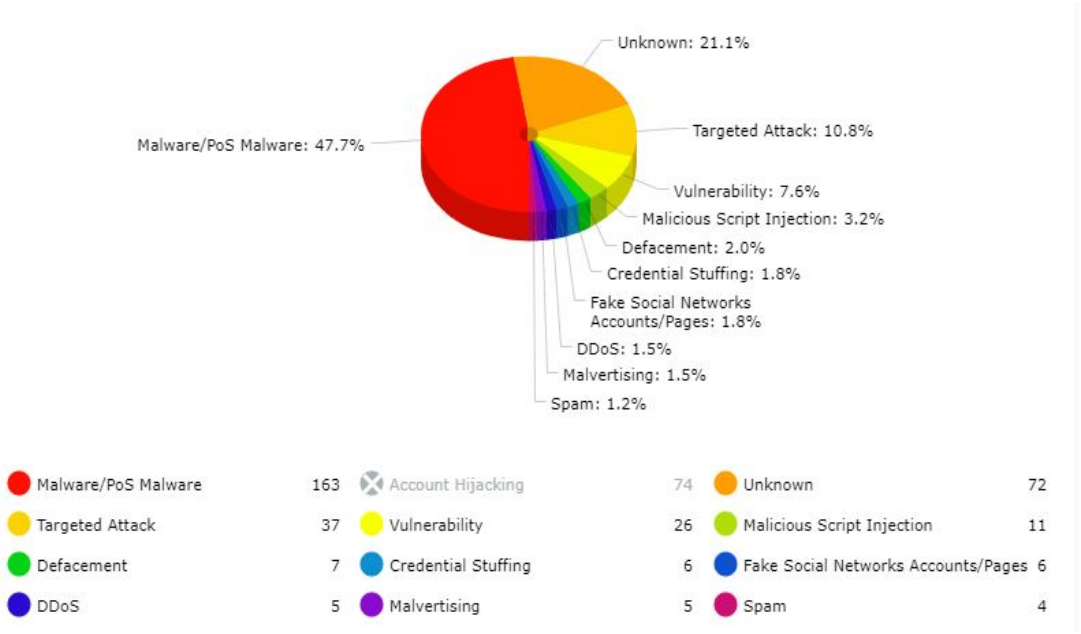

Fig. 2. Top 10 Attack Distribution for 2018 [16]

I. Literature Overview

This literature review is a first step in understanding the existing paradigms and debates around AI before narrowing the focus to more specific applications and subsequently, policy-recommendations.

Agents can be defined as autonomous to solve problems in computational structures capable of efficient functioning in dynamic and open environments [2]. They are often used in environments where they communicate with the environment itself, exchange information with them, and other agents (people or software) to resolve conflicts.

Ganpathy et al. in 2012 offer a method that combines intelligent distance factor agents (IAWDBOD) and intelligent multiclass vector (IAEMSVM) algorithms. The results for DoS, Probe, and other types of attacks are 99.77%, 99.70%, and 79.72%. The main advantage of this method is that it reduces negative positive results [3].

Jain et al. offer in their article a detailed analysis of the various penetration systems using mobile agents [4].

The main advantages of intelligent agents compared to other methods of artificial intelligence can be summarized as:

As already mentioned, the use of RL in detecting penetration has not been extensively investigated, and even less so in the widespread penetration detection. Some of the earlier studies were made by Cannady [6] [7]. Cannady shows that neural networks are realistic solutions when trained for a particular problem domain with representative learning data sets. However, they cannot adapt to the new data until they are brought online and re-qualified with the new sets of representative data. To cope with this problem, it uses a neural network CMC (Cerebellar Articulation Controller Articulation Controller). This type of neural network has an opportunity for online learning. They use a three-layer feedforward mechanism designed to produce a series of I/O images. In this study, the single IDS agent learns how to detect a denial of service attack based on ICMP and UDP. The system initially learns how to detect ICMP attacks and learns how to recognize new attacks based on the UDP protocol through previous knowledge and continuous retraining.

An approach used to detect HIDS penetration is based on monitoring the sequence of system calls. These calls are emitted by a process running on the host and grouped into different traces. Each trace contains a list of system calls generated from start to finish processes. To apply machine learning techniques using system call sequences, researchers often construct a transient model using labeled examples of normal work and attack. State of the model is determined by short sequences of system calls in one track. Xu and Xie [8] have applied the hidden branded models (CMM) and RL to detect penetration by studying the probability of transitioning the state. They claim that there is uncertainty in state modeling in IDS and CMM are able to offer a suitable alternative to the problem. They use a linear function approximation technique and a temporary difference algorithm to update the value function. The system was trained off-line and triggered a reward of -1 for normal activity and +1 for attack. Authors report positive results compared to other machine learning techniques that use the same set of drills and ratings from the DARPA intrusion detection series [9].

In another work, Xu and Lou [10] apply temporal difference methods [5] to model the dynamic behavior in the HIDS approach. To approximate the value function and extract functions, they use a kernel-based LS-TD (λ) algorithm. As described by the authors, the LS-TD (λ) algorithm is a nonlinear functional evaluation that uses high-quality space for

functions and at least quadratic TD training. To evaluate their approach, authors use traces of system calls from the sendmail application. They showed positive and better results compared to the previous performance, using the methods of hidden branded models.

Miller and Inoue [11] use DIDS training (distributed IDS) training called Perceptual Intrusion Detection System with Armor (SPIDeR). The system consists of heterogeneous agents that detect intrusion and communicate through a blackboard system. All agents have a three-layer architecture, consisting of sign-on detection for known infiltrations, SOM anomaly detection array, and a third layer of information to be collected for further analysis. Remote agents perform intrusion detection and send their voices through the blackboard system to the sense of local activity. Through the RL process, the central system calculates and, in turn, rewards the agents according to their effectiveness. Authors evaluate SPIDeR using a KDDCup'99 dataset with positive results.

Through CMM, RL, and behavioral analysis of IP addresses, Xiu et al. [12] offers DIDS focused on detecting DoS / DDoS attacks. The architecture is composed of a group of sensors that have partial environmental visibility. Due to communication limitations, sensors cannot send all their sensory information. Instead, they learn to recognize local attacks and transmit them to a central server. Although authors report high levels of detection, a possible drawback to this approach is the use of an easily accessible source of information (IPs).

Awerbuch, Holmer and Rubens [13] apply RL for security and routing. They have developed a secure routing architecture for wireless routing. The system is built up by a group of routers that share communication via secure channels. For packet routing, they use RL and are able to recognize DoS attacks against the routing infrastructure.

II. Experimental Setting

Dynamic environments have not only a multiple agents but also a number of consistent solutions that increase their complexity. In these settings, agents must coordinate and discuss the current state of their dynamic environment with very limited information. Typically, agents in a dynamic environment cannot monitor the actions of other agents or see what remuneration they receive as a consequence, although the actions of other agents affect their immediate environment, along with the rewards they can get. In very complex environments agents may not be aware that other agents exist and can interpret their environment as non-stationary. Similar, equally complex environments allow agents to access information, but value spaces are not conducive to learning because of their complexity and the necessary coordination between agents. Before an effective multi-agent approach can be developed, all these challenges need to be addressed. A standard learning model to strengthen multiple agents is presented in Figure 2.

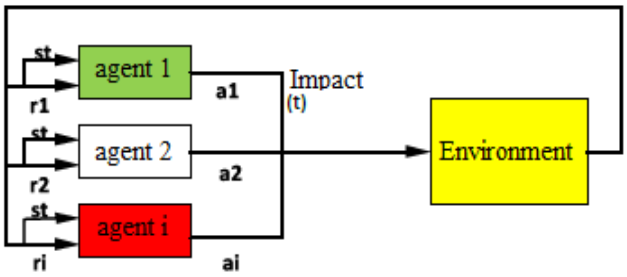

Fig. 2. Multiple agents acting in the same environment

As a result of the search for adaptive multi-agent systems and the complexity of dealing with interacting learners, an increasing number of researchers have worked to develop reinforcement methods. This field of learning uses research to reinforce learning that takes place within the framework of AI.

The IDPS's cohesive intelligence complexity includes three main concepts for discovery management: Strength Training, Knowledge Management (KM), and Multi Agent Management (MA) in the core architectural project. Collaboration between optimization techniques describes the clear idea of collaborative learning based learning to meet Hybrid Intelligent IDPS requirements.

The conceptual model of the computer network protection system is based on three layers. The first layer shows traditional system components that monitor and collect audit data through sensors, analyze data and detect intrusions, generate alarms, and declare the correct response through trigger devices. The next two layers are advanced layers based on the proposed Hybrid Intelligent IDPS.

The second layer shows traditional AI techniques that communicate with each other for more accurate results. Knowledge management makes it possible to characterize the anomaly profile knowledge as a set of related concepts in the anomaly domain. Multi-agent manager policies are used to predict abnormal behavior. They are combined with adaptive optimization techniques such as reinforcement to detect penetration and feed the results into components including self-optimization and self-learning. The latter components of the layer are defined in the principles of real-time autonomous calculation without human intervention. This removes the human factor as a cause of error.

Despite the availability of several standardized datasets, there is no standardized methodology for IDS. Today, the customized methodology is predominant among intrusion detection researchers.

Normally, a simulation environment is used to evaluate IDS in complex environments involving large networks. Background traffic and attack patterns can be injected by pre-compiled kits or by creating specific templates in accordance with the evaluation requirements.

Although IDS assessment through the test network seems to be the best method, it is not always possible due to time and resources constraints. More specifically, in our case, it was very difficult for us to use this method to evaluate DIDS. The study described here aims to develop an intrusion detection system capable of recognizing and categorizing DoS and DDoS attacks. This architecture compiles a large number of

sensor agents located in a variety of locations on a computer network. To build this network, we needed costly resources like network devices and links that were not available for this research. For this reason, it is impossible to build a rating network made up of real devices. Instead, a simulation environment was selected for evaluation of the studies. In summary, it is used in a simulation environment instead of a rating network composed of real devices and real traffic for the following reasons:

- Speed: Simulated environment tests will be faster than those in a real network. Tests in a simulator can be performed for part of the time that would be required in the test network. For example, tests that simulate five minutes of network activity can be processed by the simulator in just a few seconds. In addition, starting a test in the system may take several minutes or hours only due to the setup process. In the network simulator this process is done by scripts and takes almost no time.
- Resource Restrictions: In order to test the scalability and the ability to learn the system to its maximum capacity, you will need to add a large number of agents with different capabilities, several cells and hierarchical levels. In addition, in order to test system interoperability in intersectional domains (for example, to simulate interconnections between Internet domains), we will need several SAs and DAs. In both cases, we do not have the necessary hardware to build these complex networks.

To effectively evaluate our approach using a network simulator, we require it to have specific features that minimize differences with a real-network evaluation. The simulation environment must meet at least the following criteria:

1. Must provide a realistic model of at least: end nodes (hosts), routers, data connections, waiting methods, delays, lost packets, and the TCP / IP stack.
2. Must provide emulation of network applications and protocols such as FTP, HTTP, VoIP, etc.
3. It should also provide an interface to add source code to implement new capabilities for current objects
4. We need to be able to inject pre-recorded network traffic for future extensions of this study.
5. You must maintain a large number of agents and links to various topologies such as bus, star, ring, etc.

In order to carry out the experiments at the various evaluation stages, this study uses open-source simulation software, namely the NS2 network simulator version 2 (NS2). The reason for using the app is its openness and publicity, as well as free use

III. Results

To assess the IDS, for each possible test value, there are two types of errors: false positive (FP) and false negative (FN). FP arises when an event is predictable as compulsive, but it is actually normal until FN happens when a really intrusive event happens without being recognized as one. On the other hand, the true positive (TP) measures the share of actual positive results that are correctly identified as such, while the true negative (TN) measures the proportion of negatives that are correctly identified as such. The effectiveness of each classifier can be quantitated using the Measurement of Detection (DR) and Total Accuracy (OA) measures. DR shows the percentage of true breakthroughs that have been successfully detected:

$$DR=\frac{TP}{TP+FN}\times 100\% \quad (1)$$

OA is calculated as the total number of correctly classified infiltrations divided by the total number of observations:

$$OA=\frac{TP+TN}{TP+TN+FP+FN}\times 100\% \quad (2)$$

Effective IDS requires a high level of DR and OA, while maintaining low levels of false alarms. Accuracy is critical to developing an effective IDS, as the high speed of FP or low DR will make it virtually unusable.

Scenarios for testing were developed to verify and evaluate the effectiveness of architecture.

Simulations for detecting DDoS attacks were performed and the results are summarized in Table 2.

Table 2. Hybrid model results

ATACK (%)	MAS			MAS+RL+FL		
	FP%	FN%	OA	FP%	FN%	OA
1	1.20	1.10	57.78	1.20	1.10	60.10
5	1.40	1.30	58.20	1.30	1.20	60.05
10	1.90	1.70	58.28	1.50	1.60	60.13
15	2.10	2.00	58.68	1.70	1.80	60.70
20	2.40	2.20	58.33	1.90	2.00	62.03
25	2.60	2.30	58.25	2.10	2.30	61.83
30	2.80	2.60	58.95	2.40	2.50	62.38
35	2.90	2.70	59.30	2.60	2.70	62.58
40	3.20	3.00	59.58	3.10	3.00	62.68
45	3.40	3.20	59.78	3.20	3.40	62.70
50	3.90	3.50	59.80	3.30	3.70	63.65
55	4.10	4.00	59.70	3.50	3.80	65.08
60	4.50	4.30	60.50	3.70	3.90	66.05
Average	2.80	2.60	59.01	2.42	2.54	62.30

Obviously, the modified mechanism achieves the greatest increase in detection accuracy. From figure 3 it can also be concluded that the accuracy of detecting a percentage of the attack is higher than the other methods.

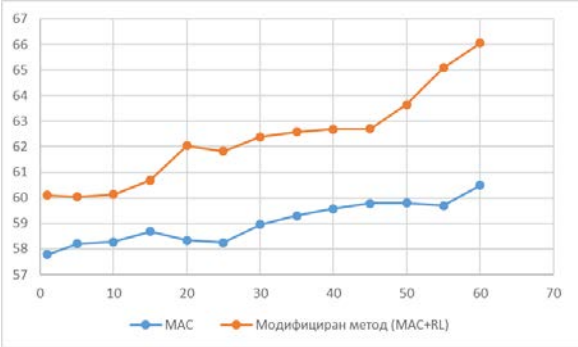

Fig. 3. Comparison of detection accuracy values

IV. Conclusion

Agent-based technology as a concept is used in almost every field and opens up opportunities to find effective solutions when the application area is a mix of different issues. Exploring and using agent-based technologies in detection and prevention systems is a complex and up-to-date task.

Scenarios for testing have been developed to verify and evaluate the effective-ness of the proposed hybrid intelligent system.

Experiments have been successfully conducted to verify and evaluate individual components and the entire platform. These experiments verify that the system meets the requirements of the specifications.

Due to the fact that the study is based on simulations and has a dependency on training data, convergence is not guaranteed.

Future work will include exploring more artificial intelligence methods combined with others to reinforce the effectiveness of detecting attacks on information security. Different results will be examined and presented and material will be produced to show in which cases what methods and means of artificial intelligence are appropriate to combine. Another task that is not widespread in penetration detection systems is the application of validation training and especially its combination of multi-agents and fuzzy logic

Acknowledgment

This research is conducted and funded in relation to the execution of a scientific-research project № H07/56 “Increasing the level of network and information security using intelligent methods” under the contract with National Science Fund in Bulgaria.

References

[1] K. C. 1999, „Computer network intrusion detection webpage on SIGKDD website,“ 199. [Online]. Available: http://www.sigkdd.org/kdd-cup-1999-computer-network-intrusion-detection.

[2] M. Luck, P. McBurney и C. Preist, Agent Technology: Next Generation Computing, AgentLink II, 2003.

[3] S. Ganapathy, P. Yogesh и A. Kannan, „Intelligent Agent-Based Intrusion Detection System Using Enhanced Multiclass SVM,“ Computational Intelligence and Neuroscience, 2012.

[4] C. Jain и A. K. Saxena, „General Study of Mobile Agent Based Intrusion Detection System (IDS),“ Journal of Computer and Communications, book 4, pp. 93-98, 2016.

[5] R. S. e. a. Sutton, „Reinforcement learning: An introduction,“ Cambridge Univ Press, 1998.

[6] C. J., „Applaying CMAC-based on-line learning to intrusion detection“, Proceedings of the International Joint Conference on Neural Networks, 2000.

[7] C. J., „Next Generation Intrusion Detection: Autonomous Reinforcement Learning of Network Attack“, Proceedings 23rd National Information Systems Security Conference , 2000.

[8] X. X. и X. T., „A Reinforcement Learning Approach for Host-Based Intrusion Detection Using Sequences of System Calls“, Proceedings of the International Conference on Intelligent Computing, 2005.

[9] J. H. D. F. J. K. K. D. R. Lippmann, „The 1999 DARPA off-line intrusion detection evaluation,“ Computer Networks, том 34, № 4, pp. 579-595, 2000.

[10] X. X. и L. Y., „Kernel-Based Reinforcement Learning Approach to Dynamic Behavior Modeling of Intrusion Detection“, Lecture Notes in Computer Science, 2007.

[11] A. I. P. Miller, „Collaborative intrusion detection system. In North American Fuzzy Information Processing Society“, 22nd International Conference of the NAFIPS, 2003.

[12] Y. S. Z. H. X. Xu, „Defending DDoS Attacks Using Hidden Markov Models and Cooperative Reinforcement Learning“, Lecture Notes in Computer Science, 2007.

[13] D. H. H. R. B. Awerbuch, „Provably Secure Competitive Routing against Proactive Byzantine Adversaries via Reinforcement Learning,“ John Hopkins University, 2003.

[14] R. Trifonov, S. Manolov, R. Yoshinov, G. Tsochev, G. Pavlova, An adequate response to new Cyber Security challenges through Artificial Intelligence methods. Applications in Business and Economics., WSEAS TRANSACTIONS on BUSINESS and ECONOMICS, Volume 14, 2017, E-ISSN: 2224-2899

[15] G. Tsochev, R. Trifonov, G. Popov, A security model based on multi agent systems, participation in XXIX International Conference on Information Technologies InfoTech-2016, St. St. Constantin and Elena resort, Bulgaria 2016.

[16] https://www.hackmageddon.com/2019/05/23/q1-2019-cyber-attacks-timeline/