Intellectualization of information and communication systems vulnerabilities validation process

Kyrychok Roman

State University of Telecommunications, 7 Solomenska street, Kiev UA-03110, Ukraine, kyrychokr@gmail.com

Abstract. *The paper proposes a new approach to the intellectualization of information and communication systems vulnerabilities validation process during the active analysis of their security, the interconnection of the tasks of validating vulnerabilities, namely the tasks of verifying and confirming the possibility of implementing detected vulnerabilities through exploits and delivering the corresponding payload, with reinforcement learning is established.*

Keywords: *information and communication system; security analysis; validation of vulnerabilities; exploit, reinforcement learning.*

I. Introduction and Problem statement

Currently, one of the most common vectors of attack remains cyberattacks using software and hardware vulnerabilities. Their implementation is possible mainly through certain "operational" gaps that arise during the operation of information and communication systems as a result of administrative errors or untimely software updates or installation of additional patches, moreover, in the absence of a regular audit of information security, vulnerabilities may remain "uncovered" for years. Along with this, the threshold for entering the cybercriminal segment is reduced due to the automation of vulnerability exploitation tools, the availability of open databases, that are almost ready for use, exploits (Exploit Database, Inj3ct0r and others) and even entire exploit packs (Magnitude, Underminer, Purple Fox and others), which can easily be found and purchased on the darknet and conduct with them full-fledged cyberattacks on the infrastructure of target organizations.

Under such conditions, the use of preventive security methods, including active security analysis, remains promising, allowing not only to identify vulnerabilities but also to validate them, i.e. to confirm that a particular vulnerability can be realized, thus establishing an actual level of information systems and networks security.

To minimize the main drawbacks of active security analysis, namely, reducing the requirements for the qualification of experts and routine analysis itself, which is especially important for large networks such as corporate networks, where may be thousands of vulnerabilities, resort to automation and intellectualization of the validation process of found vulnerabilities. However, after analyzing these approaches [1-3] it should be noted that their effectiveness remains low, because:

- automation occurs mainly due to the sequential verification of vulnerabilities by the means of exploitation, i.e. through sequential launching of all selected exploits, taking into account simple criteria (operating system family, service, exploit rank and others). At the same time, it should be noted that most of them do not work, which indicates that the decision to use the selected exploit is false; moreover, the implementation of an incorrectly selected exploit in general can lead to complete failure of the target system;
- intellectualization is carried out through the use of classical methods of machine learning (training with and without a teacher), while leaving open the issue of obtaining quality data for training such systems.

The most promising solution to these problems may be the use of the reinforcement learning. Since the reinforcement learning itself is used in cases the machine needs to correctly perform the tasks assigned to it in the external environment, having many possible options for action and the ability to interact with this environment in real time.

II. Problem solution and Results

The reinforcement learning was developed in the works of R. Sutton and E. Barto [4] based on the theory of adaptive behavior developed by M.L. Tsetlin [5]. At the same time, it should be noted that in the method of reinforcement learning, concepts such as agent, environment, and reward are introduced, that directly describe the process of optimization of a certain task. The general scheme of the reinforcement learning process is shown in Fig. 1., which shows the interaction of the agent with the environment in discrete moments of time $t = 0,1,2\ldots T$, which are also called steps. The agent is some autonomous system, which has the ability to obtain information about the state of the environment (situation) and to affect through certain actions, which lead to changes in the situation. This means that the environment is an object or everything outside of the agent with what it interacts.

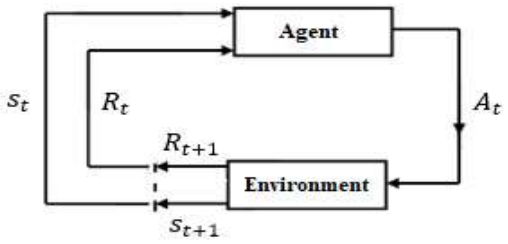

Figure 1. The general scheme of the reinforcement learning process [4]

At each time step t the agent gets a certain view of the state of environment $S_t \in S$, where S – is the final set of all possible states, and on the basis of this view selects the action $A_t \in A(S_t)$, where $A(S_t)$ – the finite set of actions that are available to the agent in the state S_t. In the next step, as a result of its action, with the help of evaluative feedback, the agent receives numerical reinforcement $R_{t+1} \in R \subset \mathbb{R}$, which can be both positive, $R_t > 0$ (reward), and negative, $R_t < 0$ (penalty), on the basis of which it forms a certain idea about the optimality of the choice made and finds itself in a new state S_{t+1}.

Thus, the agent that is learning has no input data about the need to perform a specific, predetermined "correct" action at a certain stage, moreover, it is often assumed that it does not even have any initial idea of the properties of the environment with which it interacts. On the other hand, the agent is able to make its own decisions about the choice of an action, by trial and error, to obtain reinforcement values, evaluating the performed actions and gradually improving its knowledge about the environment with which it interacts.

The main interconnection between the tasks of vulnerability validation, namely, the tasks of verification and confirmation of the possibility of implementing the discovered vulnerabilities through the use of exploits and delivery of the corresponding payload with the theory of reinforcement learning can be expressed as follows (Fig. 2):

- the vulnerability exploitation (validation) tool ↔ the agent;
- selected exploits of target information system vulnerabilities ↔ a set of action A ;
- the validation tool chooses a vulnerability exploit and implements it ↔ the agent chooses and performs a certain action;
- the validation tool received the result of an attempted operation by revoking the target system ↔ the agent received numerical reinforcement for the performed action.

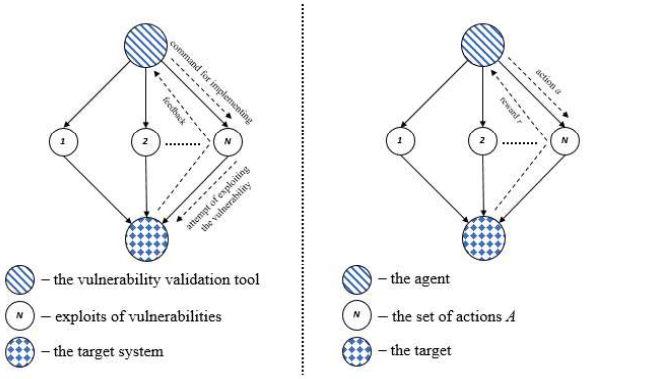

Figure 2. Relationship between the Vulnerability Validation tasks (left) and the Training task with Reinforcement (right)

Thus, it follows from the above that during the intelligent validation of vulnerabilities, exploits are ordered and implemented. During this process, the validation tool chooses the next exploit from the list of available ones (i.e. theoretically corresponding to the target system) and, by default, tries to execute it, waiting for the response from the target system and numerical reinforcement for the selected exploit. Based on this, it evaluates the optimal decisions made to use a particular exploit.

III. CONCLUSIONS

The proposed approach to the intellectualization of the vulnerabilities validation process of information and communication systems based on the use of the reinforcement learning will optimize the sequence of exploitation of likely vulnerabilities of software and hardware platforms in the target system, as well as reduce the percentage of false decisions on the use of selected exploits.

REFERENCES

[1] J. Luan, J. Wang, M. Xue, “Automated Vulnerability Modeling and Verification for Penetration Testing Using Petri Nets”, ICCCS (2), pp. 71-82, 2016.

[2] D. Wu, Y.-F. Lian, K. Chen, Y.-L. Liu, “A security threats identification and analysis method based on attack graph”, Jisuanji Xuebao (Chinese Journal of Computers), vol. 35, n. 9, pp. 1938–1950, 2012.

[3] C. Sarraute, “Automated attack planning”, Ph.D.thesis, School of Engineering, Buenos Aires, Argentina, July 2nd, 2012.

[4] R.S. Sutton, A.G. Barto, “Reinforcement Learning: An Introduction second edition”, The MIT Press, Cambridge, MA, 2018.

[5] M. L. Tsetlin, “Automaton Theory and Modeling of Biological Systems”, Academic Press, New York, 1973.