UDC 621.391

O.S. SAVENKO, S.M. LYSENKO, A.F. KRYSHCHUK

Khmelnytsky National University, Khmelnytsky, Ukraine

MODEL OF THE COMPUTER SYSTEM DIAGNOSIS PROCESS FOR BOTNET PRESENCE IN CORPORATE AREA NETWORK

The new model of the computer system diagnosing process for botnet presence in the corporate area network is proposed. It is based on the use of multi-agent system. The new botnet detection technique based on multi-agent system with the use of fuzzy logic is proposed. The detection is performed in the situations of priori uncertainty of the botnet presence in the corporate area network with taking into account the botnet demonstrations in the several computer systems available in the network. Fuzzy expert system for making conclusion about botnet presence degree in computer systems is developed.

***Key words**: bot, botnet, antiviral diagnosis, model of antiviral diagnosis process, antiviral agent model, multi-agent system model, fuzzy logic, expert system.*

Introduction

The most numerous and danger malware during the last years is a new malware class – botnet, that is the cooperation of Trojans and worm-viruses.

They are the main base for such danger acts as distributed denial of service attacks, malware distribution, phishing, theft of confidential corporate data, organization of anonymous proxy servers etc. The peculiarity of botnet is the using of specialized commands and controlled channels of interaction that provides the updating of functional bots' parts of and actions features. Some botnet performs some illicit monetary activities [1-2].

1. Related works

A lot of models and techniques for botnet detection based on them have been developed in recent years. These methods can be categorized into honeynet-based methods and based on passive traffic monitoring.

Honeynet-based Methods. Honeynet is a powerful tool for understanding botnet technology and characteristics, and tracking botnet behaviors. It is not very effective in botnet disruption.

Passive Traffic Monitoring. Another approach is setting up vantage points to passively monitor the real Internet traffic and to detect or extract the botnet related packets [3].

Behavior-based Detection: behavior based detection methods can be further categorized as signature based and anomaly based.

Signature-based Detection: a major weakness of the signature based detections is that they are limited to detect only the known botnets.

Anomaly-based Detection: this algorithm does not require prior knowledge of a botnet and has low false positive and false negative rates.

DNS-based Detection: a hybrid of behavior based and data-mining based techniques performed on DNS traffic. The main drawback of this approach is the high processing time required for detailed monitoring of the huge scale of network traffic [4].

That's why the actual task is the development of a new model of the computer system (CS) diagnosing process for the botnet presence in the corporate area network that enables the construction of more perfect techniques for new botnet detection.

2. Model of the process of the computer system diagnosing the botnet presence in the corporate area network

The object of the research is the computer systems diagnosing process for the bot detection as part of a botnet. That is why an important task is the development of diagnosing process model, that should include the ability to display features of the process and use the of botnet model described in [5].

The antiviral diagnosing process for botnet presence in the computer system that belongs to the corporate area network (CAN) is performed. The diagnosis process is based on multi-agent system (MAS). Each agent of MAS includes a set of sensors that perform antiviral diagnosis.

Let us divide the diagnosing process into four subprocesses: the monitoring of events in each computer system of the corporate area network; the computer system scanning for malware; implementation of the communication between agents of the multi-agent system; the processing of the information received from the sensors in order to make the conclusion about botnet presence in CAN.

The CS monitoring is performed since the computer system has been launched. The scanning procedure of computer system is performed on user's demand or in a specified time.

Let us present the diagnosis process model as a set

$$D =< \beta, \psi, \sigma, \theta >, \quad (1)$$

where β - monitoring process, ψ - computer system scanning process, σ – communication process, θ - processing of information from sensors with further implementation the conclusion about the possible bot presence in computer system.

In this case the formalized scheme of the diagnosing process of CS for the bot presence of can be presented in fig.1.

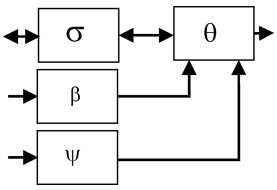

Fig. 1. The formalized scheme of the computer system diagnosing process for the bot presence

2.1 Model of the multi-agent system for computer system antiviral diagnosis

Multi-agent system contains the set of agents that are used for antiviral diagnosis. Also multi-agent system performs the communication functions for the agent actions coordination in the present moment, and for its cooperation (information exchange).

In order to achieve its goals agents are united into the in groups and work together by collecting and sharing their knowledge and capabilities to each other.

Let us present the multi-agent model as a set

$$B =< H, A, L, O >, \quad (2)$$

where H – the set of the corporate network activities; $A = \{A_1, ..., A_i\}$ - the set of agents as part of the antiviral multi-agent system; $L = \{L_1, ..., L_i\}$ - the set of agent actions; $O : H \times L_1, ..., L_i \to H$ - function that describes the agents' actions for possible network's reaction.

2.2. Agent model

Let us consider the agent as some system that is functioning in the computer system and is a part of the antiviral MAS. Agent is able to interact with other agents to make autonomous rational action to achieve some goals.

Taking into account the agent functionality let us submit the agent model as the set

$$A =< P, I, L, C, \rho, \mu >, \quad (3)$$

where P – processor, that makes a conclusion about the possibility of computer system infection with the bot based on the received data and its knowledge; I – a set of the agent states; L – a set of the agent actions, $L = \{L_1, ..., L_i\}$, where $L_i = \{l_1, ..., l_i\}$ – set of the effectors actions, that affect the diagnosis objects; $C =< Z, T, R, V >$ - communication unit that executes the information exchange between agents, where Z - system information, T - a set of agent results A_i (information, that is sent from other agents), R - processor result (information, that is to be sent to other agents), V - the function describes which signals the agent A_i will send to each agents in current time, V_{A_i} : $Z \times R \cap Z \times T \to V$.

The processor can be presented as

$$P =< U, W, R >, \quad (4)$$

where $U =< R_{S_i}, T, R >$ - agent memory which contains R_{S_i} – a set of results produced by sensors; T – a set of agent results, R - a set of results produced by processor which indicates the possible computer system infection with a bot;

$W = \{X, Y\}$ – a set of the rules for the making of the botnet presence in CAN, and the knowledge about the possible botnet demonstrations (activities) in computer systems the belong to CAN.

The set of the agent states can be represented as the set

$$I =< S_i, E, R_{S_i} >, \quad (5)$$

where S_i - a set of the agent sensors [6, 7]; E – a set of the diagnosis objects.

$\rho : I \times R_{S_i} \cap I \times T \to I$ - update function of the agent stage that takes into account the sensors information and computer network information;

$\mu : I \to L$ - the decision function that associates the current internal agent states with some action.

2.3. Agents communication

Communication is used by agents to coordinate their actions in current moment. Agents need to share information that may affect the agents. Agents need to share information as quick as it possible in order to respond to changes in the corporate area network.

One agent can send to each agent the only one signal at the moment. Function V_{A_i} is changing during the interaction with the CAN, and the sent or received signals affect the decision of other agents' actions. The choice of action and state change is performed in two stages: defining and sending signals, and the choice of the action.

Let us install an antiviral agent into each computer system (fig. 2, a). For example, an agent A_2 has obtained data (results) from sensors R_{S3}, R_{S4} and R_{S6}.

For instance, the degree of suspicion has the value that overcomes the pre-set level of the danger *m* (fig. 2, b). Using the communication unit C of the agent sent the messages are sent to other active agents of MAS (fig. 2,c). Agents store the message and return the processed information (fig. 2, d). After receiving messages from other agents processor P of the agent A_2 concludes the further action towards the suspicious object.

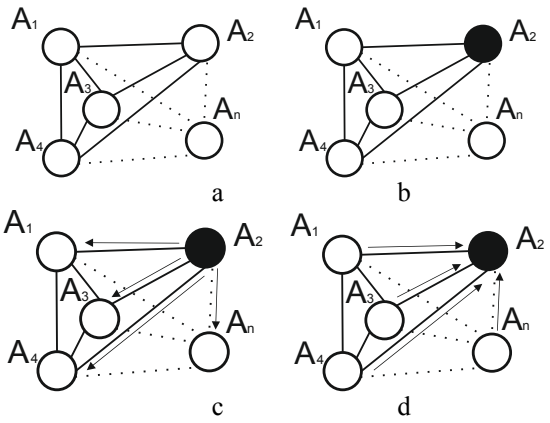

Fig. 2. The communication between the agents

3. Multi-agent based technique for botnet detection in computer systems

3.1 Antiviral agent of multi-agent system

The new techniques for the botnet detection based on proposed the model of diagnosing process is proposed. For this purpose we have to construct of a schematic map of connections which is formed by corresponding records in each antiviral agent of multi-agent systems for some corporate area network. All agents based on this information communicate with each other. Botnet detection process can be presented as a scheme shown in fig.3.

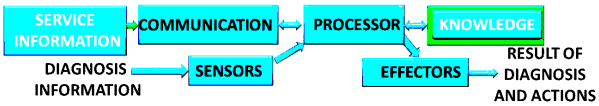

Fig. 3. The scheme of antiviral agent of multi-agent system

A new technique for determining the degree of presence of botnet is proposed. Offered method is based on analyzing of the bots actions demonstration in situations of intentional change of connection type of probably infected CS. This approach is performed in the case of insufficient (low) values of suspicion software, but this suspicion is present in a definite amount of CSs of the corporate area network.

During computer system functioning the antivirus detection via sensors available in an each agent is performed. The antivirus diagnosis results are analyzed in order to define which of sensors have triggered and what suspicion degree it has produced. If triggering sensors are signature S_1 or checksum S_2 analyzers or API sensor S_5, the results R_{S1}, R_{S2} or R_{S5} are interpreted as a 100% malware detection. In this situation, the blocking of software implementation and its subsequent removal are performed.

For situations when the sensors of heuristic S_3 and behavioral S_4 analyzers or virtual bait S_6 have triggered, the suspicion degrees R_{S3}, R_{S4} R_{S6} are analyzed, and in the case of overcoming of the defined certain threshold n, $n \le \max(R_{S3}, R_{S4}, R_{S6}) \le 100$, the blocking of software implementation and its subsequent removal are performed. If the specified threshold hasn't overcome the results R_{S3}, R_{S4} are analyzed whether they belong to range $m \le \max(R_{S3}, R_{S4}, R_{S6}) < n$ in order to make the final decision about malware presence in CS. If the value is $\max(R_{S3}, R_{S4}, R_{S6}) < m$ than the new antivirus results from sensors are expected. In all cases the antiviral agents' information of infection or suspicion software behavior in CS must be sent out to other agents.

The main topic of this approach is to research the situation where the results of antivirus detection belong to range $m \le \max(R_{S3}, R_{S4}, R_{S6}) < n$. In this case, the antiviral agent of CS asks other agents in the corporate area network about the similarity of suspicion behavior of some software that is similar to the botnet. If the interrogated agent receives information from one or more agents about the similar of software suspicious behavior, then the probably infected computer systems are marked and map reconstruction is implemented (Fig.4). From the set of "marked" computer systems some CS is must be chosen for the changing of network connection type (reconnection) - specific network settings that prevent the network functioning of the bot in the computer system (DNS change, non-standard port connection to network, etc). The means of choosing the one CS from the "marked" is the expert system. It contains a set of rules that are present in the knowledge of each antiviral agent. This CS must meet the defined criteria.

In order to choose some CS we must analyze the features and properties of probably infected computer systems with botnet. For this purpose let take the concept of "suitability" of some computer system. Thus, we are interested in the computer system with the most relevant antivirus databases, with the highest uptime duration, with the lowest vulnerability degree of the operating system and the best result of virus diagnosis. Determination of computer system "suitability" is performed with the use of a fuzzy inference system which is present in the agent structure. Each agent of probably in-

fected CS calculates the rate of its "suitability" and then communicates with other agents in order to choose CS as the most "suitable" one for the changing the type of network connection.

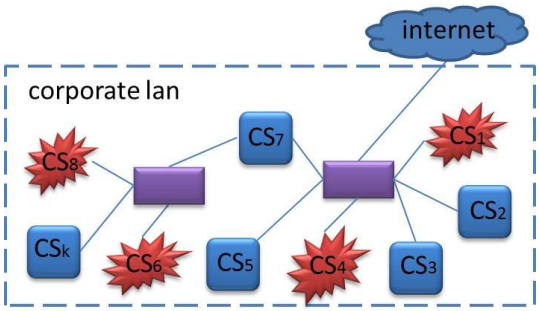

Fig.4. Marked computer systems in the corporate area network

After the reconnection of the chosen CS, the analysis of botnet demonstrations on reconnected computer system, on "marked" computer systems and other computer systems of the corporate area network and the definition of the degree of a new botnet presence in the network must be determined.

Determination of the presence of botnet network is possible due to the fact that when we change the type of connection of some computer system, bots can demonstrate itself in some way (bots can try to communicate with other elements of botnet, update lists of active bots, reconfigure itself taking to account the new lists, etc.).

Note. We must pay attention to computer system place in the topology of the corporate area network. If the computer system is a unifying node with neighboring computer systems in corporate area network (e.g. CS7 fig. 4), which can be a server or a firewall, we cannot not change the type connection of this CS.

3.2. The determination of the botnet presence degree

For the determination of the botnet presence degree in CS we must analyze botnet's demonstrations when some CS was reconnected. For this purpose all demonstrations are divided into three categories and the degrees, each of them must be determined: demonstration degree of reconnected CS, demonstration degree of probably infected computer systems and demonstration degree of other computer systems belonging to the corporate area network that probably weren't infected. To determine the possibility of the botnet presence in CS, the estimation of the demonstration degree for each of the three categories is performed. Demonstrations' degrees of three categories are presented as the fuzzy linguistic variables "demonstration degree" with three terms ("low", "medium" and "high").

The task of determination of membership function for input variable "demonstration degree" of reconnected computer we will consider as the task of the ranking for each of functions of penetration ports with the set of indications of danger. The task of determination of membership function for input variables "demonstration degree" of "marked" CSs and common (not infected) computer systems are considered as the calculating the botnet demonstration degree. We must take into account the botnet action danger, the number of computer systems and where the demonstrations took place.

Let accept ω_j^i, $0 \le \omega_j^i \le 1$ - one of the signs of the demonstration, $j = \overline{1,n}$, $i = \overline{1,\gamma}$, where γ – number of botnet demonstration, k – number of computer systems in corporate area network. The estimation of each CS can be performed with the use of formula:

$$\omega^1 = \sum_{i=1}^{\gamma} \alpha_i^1 \omega_i^1 / \gamma,$$

$$\omega^1 = \sum_{i=1}^{\gamma} \alpha_i^2 \omega_i^1 / \gamma, \qquad (6)$$

$$\omega^j = \sum_{i=1}^{\gamma} \alpha_i^{\gamma} \omega_i^j / \gamma,$$

where α_i - coefficients of the danger of some demonstration, $\alpha_1 + \alpha_2 + ... + \alpha_\gamma = 1$, $0 \le \omega^i \le 1$.

Thus if we choose some threshold value for each computer system with the estimation ω^j, for example $\tau \in (0;1]$, then we can select some group g of "suspicious" computer systems if $\omega^j > \tau$. Then we calculate d_i - number of nonzero demonstrations of d_i^j in each computer system and average value ω_i with nonzero demonstrations ω_i^j. If number of nonzero demonstrations $d_i \ne 0$ then number of nonzero demonstrations is calculated with the use of formula:

$$\omega_i = \sum_{j=1}^{n} \omega_j^i / d_i, \quad d = \sum_{i=1}^{\gamma} d_i \le \gamma \cdot k. \qquad (7)$$

We have to normalize the number $\omega_i, i = \overline{1,\gamma}$, so that $\omega_1 + \omega_2 + ... + \omega_\gamma = 1$. Then general demonstration degree of botnet presence in "marked" computer systems is:

$$P_d(d_1, d_2, ..., d_\gamma) = \frac{d!}{d_1! d_2! ... d_\gamma!} \cdot \omega_1^{d_1} \cdot \omega_2^{d_2} \cdot ... \cdot \omega_\gamma^{d_\gamma}. \qquad (8)$$

Let k', $k' \le k$ – number of "marked" as infected computer systems. Then the arithmetic middling $\overline{\omega}$ of its correspondent ω^j must be calculated. After that the number P_d is determined and is interpreted as degree of botnet demonstration in "marked" computer systems.

4. Experiments

In order to verify the offered model of the antiviral diagnosing process the new software was developed and the experiments were held.

The research has been conducting for 8 months and such results have been obtained: using proposed technique the botnet detection demonstrates better results in comparison with simple local detection.

For the implementation of an experiment 60 programs with the botnet properties (Agobot, SDBot and GT-Bot) were generated. During the experiment computer systems in the network were infected only by one botnet and it was held during 24 hours. The results of the experiment in comparison with local detection are shown in table 1 and in fig. 5.

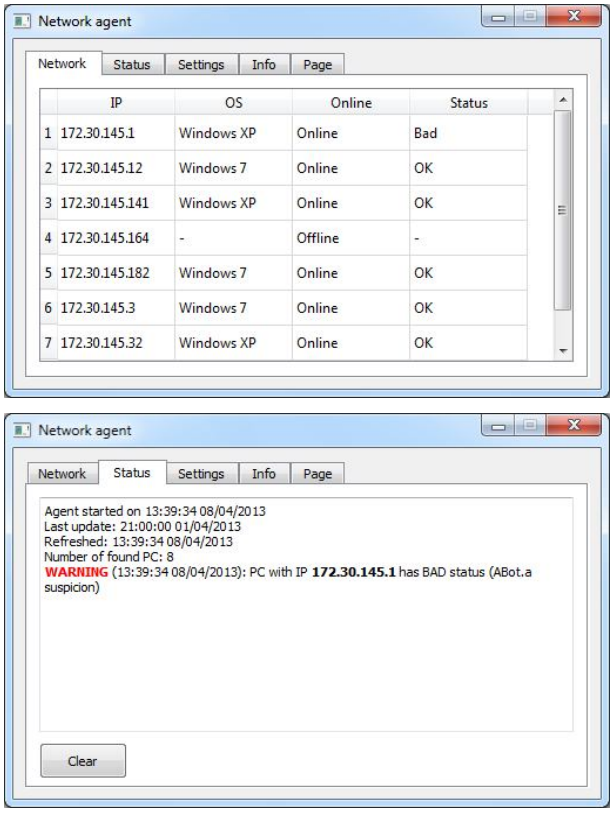

Fig. 5. The results of software

Table 1

The results of the experiments

	Local detection	MAS detection
Agobot	12	14
SDBot	18	20
GT-Bot	16	17
Overall	**46 (76,6%)**	**51 (85%)**

Experiment results prove the efficiency using the multi-agent system for botnet detection in comparison with the use of the proposed technique and without it. The increasing of the efficiency is about 7-10%.

Conclusion

The new model of the process of the computer system diagnosing the botnet presence in the corporate area network is proposed. It is based on the use of multi-agent system.

Based on proposed model the new botnet detection technique based on multi-agent system with the use of fuzzy logic is proposed. The detection is performed in the situations of priori uncertainty of the botnet presence in the corporate area network with taking into account the botnet demonstrations in the several computer systems available in the network.

With the usage of fuzzy logic, the analysis of the botnets' actions demonstrations in the situation of the intentionally computer system reconnection is performed. Fuzzy expert system for making conclusion about botnet presence degree in computer systems is developed. Fuzzy expert system takes into account the demonstration degree of reconnected computer system, demonstration degree of probably infected computer systems and demonstration degree of other computer systems available in the corporate area network that probably weren't infected.

The involvement of the developed method proves the effectiveness of the botnet detection with its growth which is about 7-10%. At the same time the increase of false positives hasn't observed. The consistency of agents in order to improve the efficiency of botnet detection is the direction of the further research.

References

1. Cooke, E. The zombie roundup: Understanding, detecting, and disrupting botnets / E. Cooke, F. Jahanian, and D. McPherson[Text] // Proceedings of the USENIX SRUTI Workshop. - 2005. – P. 39–44.

2. C. Mazzariello. IRC traffic analysis for botnet detection [Text] / C. Mazzariello // Fourth International Conference Information Assurance and Security ISIAS'08. – 2008. – P. 318 – 323.

3. Akiyama, M. A proposal of metrics for botnet detection based on its cooperative behavior [Text] / M. Akiyama, T. Kawamoto, M. Shimamura, T. Yokoyama, Y. Kadobayashi, and S. Yamaguchi // Applications and the Internet Workshops, 2007. SAINT Workshops 2007. Int. Symposium. - 2007. – P. 82–82.

4. Choi, H. Botnet detection by monitoring group activities in DNS traffic [Text] / H. Choi, H. Lee, H. Lee, and H. Kim // In proceedings of the 7th IEEE International Conference on Computer and Information Technology. IEEE Computer Society. - 2007. – P. 715–720.

5. Savenko, O. Review of botnet detection techniques [Text] / Lysenko S, Kryshchuk A // Proceedings

of the international conference "10 IEEE East-west design and test simposium, 2012. - P.479-482.

6. Savenko, O. Botnet detection based on multi-agent approach [Text] / S. Lysenko, A Kryshchuk // Radioelectronic and computer systems. – 2012. – № 5. – P. 97 – 112 (in Ukrinian).

7.Savenko, O. Multi-Agent Based Approach of Botnet Detection in Computer Systems [Text] / O. Savenko, S. Lysenko, A. Kryschuk, // 19th Conference on Computer Networks, CN 2012, Szczyrk, Poland. CCIS. - Springer, Heidelberg 2012. - V. 291. - P. 171-180.

Поступила в редакцию 14.02.2013, рассмотрена на редколлегии 13.03.2013

Рецензент: д-р техн. наук, доцент А.В. Горбенко, Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков, Украина.

МОДЕЛЬ ПРОЦЕСУ ДІАГНОСТУВАННЯ КОМП'ЮТЕРНИХ СИСТЕМ НА НАЯВНІСТЬ БОТНЕТ-МЕРЕЖ В КОРПОРАТИВНІЙ МЕРЕЖІ

О.С. Савенко, С.М. Лисенко, А.Ф. Крищук

Запропоновано нову модель процесу діагностування комп'ютерних систем на наявність ботнет-мережі в корпоративній мережі. Вона базується на використанні мультиагентних систем, використовуючи моделі ботнет-мережі. Запропоновано новий метод діагностування КС на наявність ботнет-мереж на основі мультиагентних систем з використанням нечіткої логіки. Виявлення здійснюється в ситуаціях апріорної невизначеності присутності ботнет-мережі в корпоративній мережі з урахуванням проявів «бота» в декількох комп'ютерних системах, доступних в мережі. Розроблено нечітку експертну систему для здійснення висновку про ступінь наявність ботнет-мережі в комп'ютерних системах.

Ключові слова: «бот», ботнет-мережа, антивірусне діагностування, модель процесу антивірусного діагностування, модель агента, модель мультиагентної системи, нечітка логіка, експертні системи.

МОДЕЛЬ ПРОЦЕССА ДИАГНОСТИРОВАНИЯ КОМПЬЮТЕРНЫХ СИСТЕМ НА НАЛИЧИЕ БОТНЕТ В КОРПОРАТИВНИХ СЕТЯХ

О.С. Савенко, С.Н. Лысенко, А.Ф. Крыщук

Предложена новая модель процесса диагностирования компьютерных систем на наличие ботнет-сети в корпоративной сети. Она базируется на использовании мультиагентных систем, используя модели ботнет-сети. Предложен новый метод диагностирования КС на наличие ботнет-сетей на основе мультиагентных систем с использованием нечеткой логики. Обнаружения проводится в ситуациях априорной неопределенности присутствии ботнет-сети в корпоративной сети с учетом проявлений «бота» в нескольких компьютерных систем, доступных в сети. Разработана нечеткая экспертная система для принятия заключения о степени наличие ботнет-сети в компьютерных системах.

Ключевые слова: «бот», ботнет, антивирусное диагностирование, модель процесса антивирусного диагностирования, модель агента, модель мультиагентной системы, нечеткая логика, экспертные системы.

Савенко Олег Станіславович – канд. техн. наук, доц., Хмельницький національний університет, Хмельницький, Україна, e-mail: savenko_oleg_st@ukr.net.

Лисенко Сергій Миколайович – канд. техн. наук, доц., Хмельницький національний університет, Хмельницький, Україна, e-mail: sirogyk@ukr.net.

Крищук Андрій Федорович – аспірант, Хмельницький національний університет, Хмельницький, Україна, e-mail: rtandrey@rambler.ru.