computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

ADAPTIVE INFORMATION TECHNOLOGY OF THE TROJAN' DETECTION IN COMPUTER SYSTEMS

Oleg Savenko, Sergiy Lysenko

Khmelnytskyi National University,
11 Instytutska street, 29016 Khmelnytskyi, Ukraine
e-mail: kism@beta.tup.km.ua, sirogyk@ukr.net

Abstract: *Adaptive information technology of computer systems Trojans diagnosing, which includes methods of diagnosing computer systems in monitor and scanner modes and allows improving reliability and efficiency, is developed. It is based on the behavioral model and the model of diagnosis. Computer system Trojan diagnosis software, which made it possible to detect the new Trojans with high reliability and efficiency, was developed.*

Keywords: *Trojans, computer system Trojan diagnosis, fuzzy logic, artificial immune systems, antivirus software, antiviral monitor, antiviral scanner.*

1. INTRODUCTION

The analysis of the situation of development of the malware shows dynamic growth of their quantity. Among them the special place occupies a class of viruses – Trojans which unlike virus programs penetrate into computer system (CS) for the purpose of information plunder that represents real danger [1].

The actual problem of safety of various CS is a development of new more perfect information technologies which provide increase of reliability and efficiency of anti-virus software diagnosis.

2. TROJAN BEHAVIORAL MODEL

The behavioral Trojan model which takes into account the features of Trojans and formalizes the process of functioning of Trojans in CS during its life cycle and to takes into account the destructive nature of its actions in the CS Trojan was developed:

$$M_v = \langle \Theta, S, V, L, Aff, \varepsilon, Z \rangle, \qquad (1)$$

where Θ – the set of all the Trojans; S – Trojan's life cycle stages, that are penetration, activation and executing destructive actions, $s_i \in S$, $i = \overline{1,3}$; $V = |V_{mp}|$ – the matrix of relationship in which $m = \overline{1,k}$ are functions (mechanisms) which perform ways of the penetration of Trojans to CS of the user via system ports $p = \overline{1,h}$ of network protocols; $L = |L_{ab}|$ – a matrix of relationship, in which $a = \overline{1,\sigma}$ are operations of Trojan which negatively influence the structural components $b = \overline{1,\tau}$ of operating system; Aff – a function that defines the interaction between objects of CS and Trojan, thus the set $a \in Aff(b_i, \nu_j)$ is a set of possible actions, that Trojan causes the object (objects); ε – the ratio between Trojan and its stages then for $\nu \in \Theta$ and $s \in S$, the relationship means $\nu \varepsilon s$ that Trojan is in stage s; Z – characteristic parameters of mentioned relations, $Z = \{z_k\}$ – set of destructive actions with normalized priority weights $P = \{p_k\}$ ($\sum p_k = 1$) which take into account the level of danger to the CS.

To define the relationship between the Trojans, its actions and stages of its life cycle, the designation was worked in $\longrightarrow$ that means: if $s_i \xrightarrow{a} s_{i+1}$ then the action $a \in A$ causes a transition from s_i stage to stage s_{i+1}. Then the Trojan, which has a life cycle with all stages passes a possible way, is:

$$s_0 \xrightarrow{V,L} s_1 \xrightarrow{V,L} s_2 \xrightarrow{V,L} s_3, \qquad (2)$$

where $s_i \xrightarrow{V,L} s_{i+1}$ means the possibility of customized life cycle [2].

The models of each classes of Trojan with the regard to their specificity and functional significance

and is based on its behavioral model were produced.

A process of computer systems Trojans diagnosis consists of two subprocesses of monitoring Ω, $\Omega = \{\Omega_1, \Omega_2, \Omega_3, \Omega_4, \Omega_5, \Omega_6\}$ and scanning Δ, $\Delta = \{\Delta_1, \Delta_2, \Delta_3, \Delta_4\}$. The process of Trojan diagnosis in the monitor mode consists of: Ω_1 - monitoring the flows, carried out through the system ports of the CS; Ω_2 – monitoring the execution of the system functions in the CS; Ω_3 – blocking the implementation of application functions, defined as suspicious; Ω_4 – procedure of fuzzification within the fuzzy inference system (FIS) by entering the degrees of suspicion and the degrees of computer system infection danger; Ω_5 – implementation of the fuzzy logic engine; Ω_6 – procedure of defuzzification within the FIS to determine the risk of CS infection by Trojans [3].

In order to formalize the implementation of antivirus diagnosing stages the model of the Trojan diagnosing process was developed:

$$M_\upsilon = \langle \{E, R, M_W, f_m\}, \{E, H, S, D, E_v, f_s\} \rangle, \quad (4)$$

where for steps $\Omega_1 - \Omega_6$: E – set of diagnosing objects in the monitor mode $e_k \in E$, that is the set of CS files $\Theta \in E$; R – the resultant number $R \in [0,1]$ that indicates the danger degree of infection with Trojan; ε – ratio between objects $v \in \Theta$ and stage $s \in S$; M_W – set of behavioral Trojan models; $f_m(I_m, I'_m, I''_m)$ – adaptability function of computer system diagnosis in monitor mode, whose arguments vary with the input data, where I_m – set of diagnosis information, $I_m = \langle \Theta, V, L, R, \rangle$; I'_m – set of the antivirus diagnostics results, $I'_m = \langle R_1, R_2, ..., R_n \rangle$; I''_m – set of detected malicious software, $I''_m = \langle E, R \rangle$; for steps $\Delta_1 - \Delta_4$: E – set of diagnosing objects in the scanner mode, $e_k \in E$; H – set of objects to be scanned; S – set of protected binary sequences; D – set of generated detectors, $d \in D$, E_v – set of files that were substituted by Trojan versions; $f_s(I_s, I'_s, I''_s)$ – adaptability function of computer system diagnosis in scanner mode, where I_s – set of diagnosis information, $I_s = \langle H, S, D \rangle$; I'_s - antivirus scan results represented with a set of files that were substituted by Trojan versions, $I'_s = \langle E_1, E_2, ..., E_n \rangle$; I''_s – set of updated system files or installed new software, $I''_s = \langle E'_1, E'_2, ..., E'_n \rangle$.

3. INFORMATION TECHNOLOGY OF TROJANS DIAGNOSIS

To solve this problem a new information technology (IT) of CS Trojans diagnosing was proposed. It includes Trojan behavioral models and techniques of computer systems diagnosing in monitor and scanner modes with automated setting up of antivirus diagnosis parameters with the improving reliability and efficiency. New technology allows detecting known an unknown Trojans. The proposed AIT allows making a conclusion of regarding the degree of danger of CS infection by Trojans and to reveal the fact of system files substitution by Trojans' versions.

4. TECHNIQUE OF TROJAN DIAGNOSIS PROCESS IN MONITOR AND SCANNER MODES

A new technique for computer system Trojan diagnosis in monitor mode which uses fuzzy logic and is based on behavioral model was developed. It enables to make a conclusion about the degree of danger of CS infection by Trojans. For this purpose we construct the input and output linguistic variables with names: "suspicion degree of software object" – for the input linguistic variable, and "danger degree of the infection" – for output one.

The task of determination of membership function for input variable we will consider as the task of the ranking for each of mechanisms (functions) m_i of penetration ports p_j with the set of indications of danger Z and a choice of the most possible p_j with activation of some function m_i. Then we generate a matrix of advantage $S = |s_{ij}|$. Elements of given matrix s_{ij} are positive numbers: $s_{ij} = s_i / s_j$, $0 < s_{ij} < \infty$; $s_{ji} = 1/s_{ij}$, $s_{ii} = 1$, $i, j = \overline{1,l}$, l – amount of possible results. Elements s_{ij} of matrix S are defined by calculation of values of pair advantages to each indication separately taking into account their scales $Z = \{z_k\}$; $k = \overline{1,r}$ with usage of such formula

$$s_{ij} = \sum_{k=1}^{r} s_{ij}^k \cdot p_k \Big/ \sum_{k=1}^{r} s_{jk}^k \cdot p_k . \quad (5)$$

Eigenvector $\Pi = (\pi_1, ..., \pi_m)$ is defied by using a

matrix of advantage. This eigenvector answers maximum positive radical λ of characteristic polynomial $|S-\lambda\cdot E|=0$. $S\cdot\Pi=\lambda\cdot\Pi$, where E is an identity matrix. Elements of vector Π ($\sum\pi_i=1$) are identified with an estimation of experts who consider the accepted indications of danger. The same procedure is performed for all m_i. As a result we receive a matrix of relationship $V_p=|m_i,p_j|$, in which each pair (relationship) m_i, p_j value $0\le\pi\le1$ responds. Using matrix $V_p=|m_i,p_j|$, we build matrix $V_p^*=|m_i,p_j|$ in which the relationship (m_i,p_j) is used and the elements of this relationship have value $\pi_{\max}$ ($0\le\pi_{\max}\le1$). Using matrix $V_p^*=|m_i,p_j|$, we build normalized curve for membership function $\mu_{X^p}(R)$ of an input variable. As a part of the solution of the problem the FIS using Mamdani algorithm was realized [4].

A new technique for constructing the protected sequences and generation of detectors based on the use of algorithms for artificial immune systems was produced. It makes it possible to reveal the fact of system files substitution of Trojans' versions [5].

The method involves the following steps: forming a set of files to be scanned: system libraries, executables system services and device drivers, which can be taken as the samples; generate protected sequences and detectors depending on operating system; comparson of the protected sequences with detectors at the stage of virus scanning; notification about the substitution when the protected sequences match with detector; check the suspicion of software actions.

Thus protected sequences and detectors have format for GNU / Linux operating system:

$$D_i^L=\left\langle\begin{matrix}m_1...m_i...m_{x1},u_1...u_i...u_{x2},g_1...g_i...g_{x3},\\ s_1...s_i...s_{x4},t_1...t_i...t_{x5},C_1...C_i...C_{x6}\end{matrix}\right\rangle,(6)$$

where $m_1...m_i...m_{x1}$ - file mode (type, permissions); $u_1...u_i...u_{x2}$ – identifier of the file owner; $g_1...g_i...g_{x3}$ – identifier of the group owner; $s_1...s_i...s_{x4}$ – file size; $t_1...t_i...t_{x5}$ – time of last file modification; $C_1..C_i...C_{x6}$ – CRC of the file, $i=\overline{1,n}$, n – number of detectors.

Protected sequences and detectors have format for MS Windows operating system:

$$D_i^W=\left\langle\begin{matrix}s_1...s_i...s_{z1},t_1...t_i...t_{z2},\\ a_1...a_i...a_{z3},C_1...C_i...C_{z4}\end{matrix}\right\rangle \qquad (7)$$

where $s_1...s_i...s_{z1}$ – file size; $t_1...t_i...t_{z2}$ – time of last file modification; $a_1...a_i...a_{z3}$ – file attribute (read-only, hidden, system, archived); $C_1..C_i...C_{z4}$ – CRC of the file, $i=\overline{1,n}$, n – number of detectors.

Generation of detectors is performed using the modified negative selection algorithm.

Antivirus software for computer system diagnosing based on proposed algorithms was developed. The results confirmed that the use of AIT of computer system Trojan diagnosis increases the diagnosing reliability by 5-15%, and efficiency – by 40% in comparison with the known antivirus technologies [6].

6. CONCLUSIONS

The article is devoted to solving important scientific problem – increasing reliability and efficiency of computer systems diagnosing of the Trojans existence. The new techniques for computer system Trojan diagnosis in monitor and scanner modes which allows improving reliability and efficiency, are developed. Also information technology is based on the model of the Trojan diagnosing process, which allows performing diagnosing with high reliability was developed. Computer system Trojan diagnosis software, which made it possible to detect the new Trojans with high reliability and efficiency was developed.

7. REFERENCES

[1] Yevgen Kasperskyi, *Computer hucking, 1-st edition,* Spb.: Piter, 2009. – 208 p. (in Russian)

[2] Oleg Savenko, Sergiy Lysenko, Model search process Trojans in personal computers, *Radio Electronic and Computer Systems,* 7 (2008). – pp. 87-92. (in Ukrainian)

[3] Oleg Savenko, Sergiy Lysenko, Intelligent mehod and algorithms of the Trojan search in the computers systems, *Visnyk of the Vinnytsia Politechnical Institute,* 6 (2008). – pp. 129-137. (in Ukrainian)

[4] R. Grafov, O. Savenko, S. Lysenko, Using fuzzy logic to search for Trojan software in computing systems, *Visnyk of Chernivtsi National University,* 6 (2009). – pp. 85-91. (in Ukrainian)

[5] O. Savenko, S. Lysenko. The development of process of Trojan detection using artificial immune systems, *Visnyk of Khmelnytskyi National University,* 5 (2008). – pp. 183-188. (in Ukrainian)

[6] S. Lysenko, A. Gontar, A. Shevtsov, Software development of the intelligent method of the trojan detection in personal computers, *Visnyk of Khmelnytskyi National University,* 1 (2010). – pp. 98-105. (in Ukrainian)