Methods of information systems protection

UDC 004. 732.056 **doi: 10.20998/2522-9052.2017.2.10**

S. Gavrylenko[1], V. Chelak[1], N. Bilogorskiy[2]

[1] National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine
[2] Juniper Networks, San Francisco, California, United States

INVESTIGATION OF INTRUSION IN COMPUTER SYSTEMS BASED ON THE HURST EXPONENT

The **subject** of the research in this article is the analysis of intrusion detection methods in computer systems. The purpose of the article is to develop effective methods and technologies for countering computer viruses. **Tasks**: research of modern means of antivirus protection of computer systems; a study of the Hurst index for assessing the state of the computer system; development of a software model for assessing the state of a computer system based on the Hurst index, analysis of the experimental data. The methods **used** are: self-similarity assessment of the process based on the Hurst index. The following results are **obtained**. A method for identified abnormal behavior of a computer system based on the Hurst index is proposed. It is based on the analysis of CPU and RAM. The results of the research showed that the influence of a number of viruses on the computer system leads to the aspiration of the Hurst index to an average value of 0.5, which indicates the randomness of the process. **Conclusions**. Experimental studies confirm the possibility of using the Hurst index as an integral part of the intrusion detection system in computer systems.

Keywords: computer system; unauthorized access; computer virus; malicious software; fractal analysis; Hurst exponent.

Introduction

It is difficult in our time to overestimate the importance of information safety. Information and data are the greatest value of modern society and its safety is critical not only for individual, but also for international safety. More and more data is being digitalized. With the growth of the value of information increases the demand for it, and at the same time – the number of those wishing to obtain unauthorized access to it. The easiest way to do this is through computer viruses.

For the first time in the history of Ukraine, in the summer of 2017, hacking attacks by the virus NotPetya struck banks, fuel stations, stores, websites of state structures for several hours [1]. Even sites of the Cabinet of Ministers and some of the largest media companies were paralyzed.

It should be noted that this problem is complicated by the dynamic increase in a number of mobile devices, the general switch to cloud technologies and the spread of Internet technologies. It leads to an increase in the spread of malicious software.

That is why the topic of the development of effective methods and technologies for counteracting computer viruses is relevant.

The analysis of the problem and formulation of the task

The main components of the antivirus system are usually a signature scanning module and a heuristic analysis module. However, if the scan method detects only known viruses, heuristic analyzers can detect new, yet unknown, viruses at the initial stage of their work.

The literature analysis [2–5] showed a large number of approaches and methods of heuristic analysis: intelligent subsystems based on the theory of artificial intelligence, methods of fuzzy logic, cluster analysis, the theory of neural networks, genetic algorithms, and others. The main disadvantage of the heuristic method is the high frequency of false positives.

Parametric methods based on control maps and methods of statistical data processing [6, 7] can also be used to solve the set tasks. These methods are based on the assumption that there is a template for a computer system for normal behavior and any significant deviations from it may be due to the influence of intruders. That is why it is very important to evaluate the work of the computer system with maximum accuracy.

At the same time, the more input data is analyzed, the more accurate the result of the evaluation is. At the same time, if the model or evaluation criteria are not selected correctly, parametric methods lose their basic authority, which can lead to an increase in false positives.

The conducted studies have shown that the main way of eliminating these disadvantages is to improve the information technology models and to argue the choice of criteria for assessing abnormal behavior of the computer system.

Task solution

The key parameter for fractal analysis is the Hurst index. The Hurst exponent was first used by the outstanding British hydrologist Harold Edwin Hurst when designing a dam on the Nile in Egypt to assess the inflow and outflow of water [8].

Hurst, having studied the records of the floods of the Nile for nine centuries, found regularity in this process.

He proved the possibility of distinguishing a random series from a non-random one, even if the random series is not normally distributed, linking it to the degree of self-similarity of the process. An object that has this "quality" is statistically similar in different scales - spatial or temporal, that is, it has a cyclicity.

The calculation of the Hurst exponent [9, 10] can be performed using the following formula:

$$H = \frac{\ln(R/S)}{\ln(\alpha N)}, \quad (1)$$

where H is the Hurst exponent; S is the standard deviation of a set of observations x; R is the range of the accumulated deviation Z_u; N is the number of observation periods; α is a given constant, a positive number.

$$S = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - X_a)^2}, \quad (2)$$

where X_a is the arithmetical mean of a set of observations x in N periods:

$$X_a = \frac{1}{N}\sum_{i=1}^{N} x_i . \quad (3)$$

The range of the accumulated deviation R is the most important element of the formula for calculating the Hurst exponent. In general, it is calculated as follows:

$$R = \max_{1\le u\le N}(Z_u) - \min_{1\le u\le N}(Z_u), \quad (4)$$

where Z_u is the accumulated deviation of set x from the average X_a:

$$Z_u = \sum_{i=1}^{u}(x_i - X_\alpha). \quad (5)$$

The Hurst exponent (H) characterizes the degree of self-similarity of the process as follows [11]:

1) 0 < H <0.5 is a random process that does not have self-similarity and is characterized by a tendency toward an average value;

2) H = 0.5 is a completely random process without a pronounced tendency;

3) H > 0.5 is a trend-based process that has a long memory and is self-similar.

Investigation of the Hurst exponent to assess the state of the computer system

In this paper, a study of the of the Hurst exponent, based on the analysis of the CPU and computer RAM memory for detecting the intrusion of the computer system, is implemented.

The algorithm for assessing the state of safety of the computer system is divided into four stages:

1) collecting data about the central processor and RAM usage;

2) statistical analysis;

3) estimation of the Hurst index;

4) determining the presence of anomalies.

For research, a software model has been developed that involves a variation in the number of set values. The CPU and CPU boot load values are scanned instantly and stored in the file. Information about CPU and RAM usage was recorded on 20 computers for three hours in safe mode (10,000 values) and when they were affected by various types of malicious software (VirKP55, VirMask, etc.). Received input data is fed to the input of the analysis module, which processes them and calculates the Hurst exponent.

Fig 1 shows the results of the Hurst exponent calculation for loading the central processor in safe mode. The value of the Hurst exponent goes to one, indicating the long-term dependence of the statistics.

As we can see from Fig 2, when the computer system is exposed to viruses, the Hurst exponent is characterized by direction to the average value, indicating the randomness of the process.

Identical results of the Hurst index calculation researches are obtained for the ratio of the usage factor of the central processor and RAM. In safe mode (Fig 3) we have a trend-resistant process that has a long memory and is self-similar.

When the system is affected by viruses (Fig 4), the process is completely random without a marked tendency.

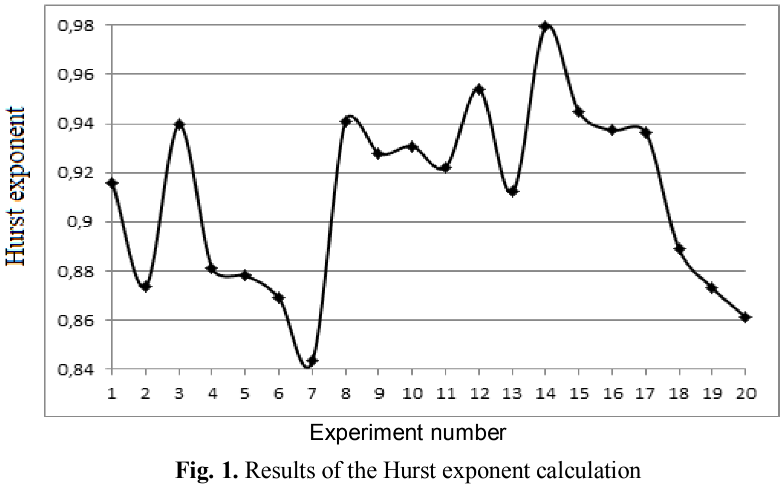

Fig. 1. Results of the Hurst exponent calculation for central processor usage in safe mode

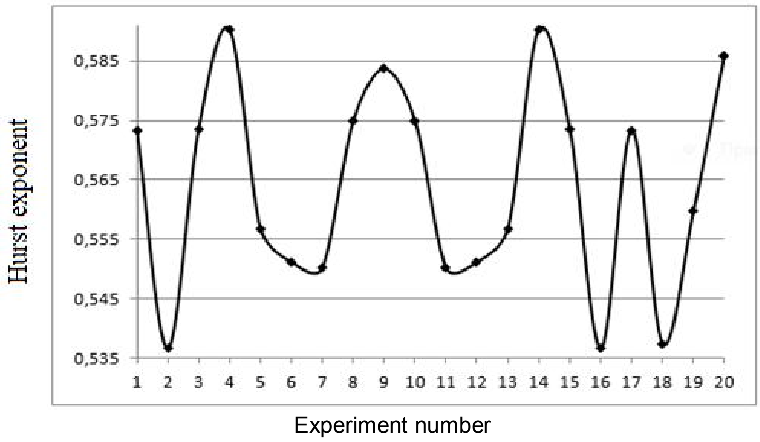

Fig 2. Research results of the Hurst exponent calculation for the central processor usage under the influence of the system by viruses

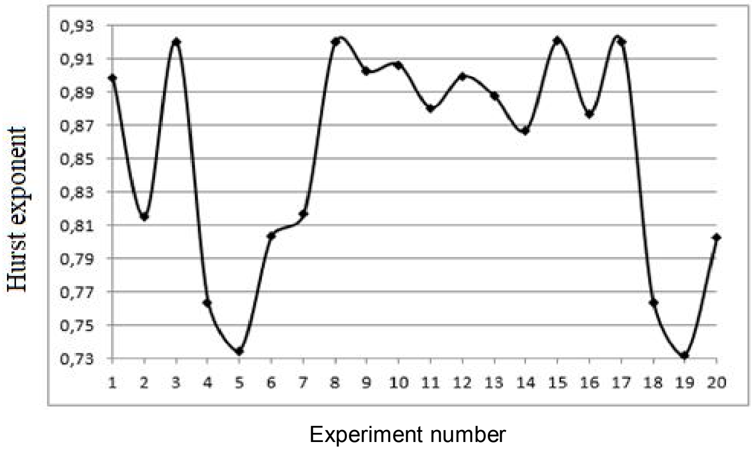

Fig. 3. The results of researches of calculation of the Hurst exponent for the ratio of the value of load of the central processor and operational memory in safe mode

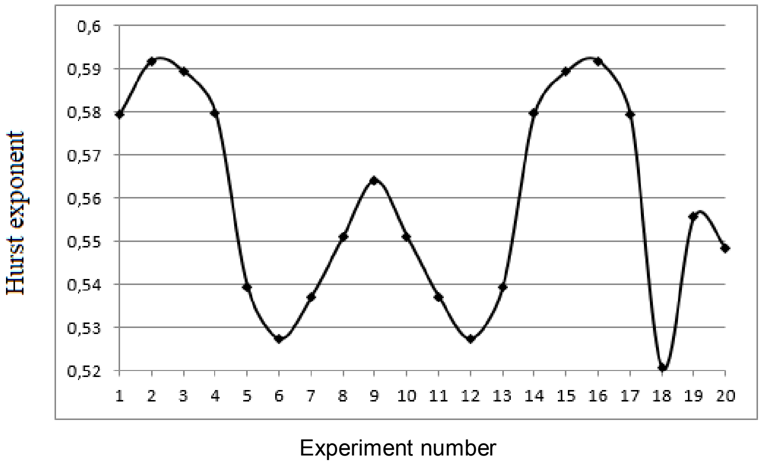

Fig 4. The results of researches of calculation of the Hurst exponent for the ratio of the central processor and operational memory usage when the system is affected by viruses

Conclusions

In this work the method of fixing anomalous behavior of computer systems on the basis of the Hurst exponent is proposed, which is based on the analysis of the central processor and RAM usage of the computer. The results of the research showed that during normal work of the computer, the value of the Hurst exponent goes to one, indicating the long-term dependence of the statistical data. The influence of a number of viruses on a computer system leads to a change in the Hurst exponent, which is characterized by a movement to an average value of 0.5, indicating the randomness of the process, which means that the process is not self-similar.

This experimental studies confirm the possibility of using the Hurst exponent as part of the system for detecting malicious software.

REFERENCES

1. *Cyberexpert estimated the damage from the Petya virus in the world*, available at: https://tsn.ua/svit/kiberekspert-ociniv-zbitki-vid-virusu-petya-a-u-sviti-953633.html (last accessed May 31, 2017).
2. Shelukhin, O.I., Sakalema, D.Zh. and Filinova, A.S. (2013), *Identifying intrusions in computer* systems, Telecom-Hotline, Moscow, 220 p.
3. Lukatsky, A.V. (2001), *Identifying attacks*, VHB-Petersburg, Sankt-Peterburg, 624 p.
4. Semyonov, S.G., Davydov, V.V. and Gavrilenko, S.Y. (2014), *Data security in computerized control systems* (monography), LAP LAMBERT ACADEMIC PUBLISHING, Germany, 236 p.
5. Semyonov S. and Gavrilenko S. (2015), "Approximating computer system operation technologies under external action through the brusselator model with perturbation in the form of dynamic chaos", *Revista RECENT*, Transilvania University of Brasov, Romania, Vol. 16, No. 1 (44).
6. Semyonov G., Gavrilenko S. and Chelak V. (2016), Developing parametrical criterion for registering abnormal behavior in computer and telecommunication systems on the basis of economic test, *Actual problems of economics*, Kyiv, Vol 4(178), pp. 451-459.
7. Gavrylenko, S., Chelak, V. and Hornostal O. (2016), "Intrusion detection in computer systems", Proceedings of the symposium "*Metrology and metrology assurance*", Sozopol, Bulgaria, 2016, pp. 342-347.
8. Hurst, H.E. (1951), Long-term Storage of Reservoirs, *Transactions of the American Society of Civil Engineers*, 116 p.
9. Shelukhin O.I. and Antonyan, A.A. (2014), "Analysis of changes in the fractal properties of telecommunications traffic caused by abnormal intrusions", *T-Comm*, No. 6, pp. 61-64.
10. Nayman, Eric (2017), *Calculating the Hurst exponent with the purpose of finding trends (persistence) in financial markets and macroeconomic indicators*, available at: http://www.wealth-lab.net/Data/Sites/1/SharedFiles/doc/forindicators/articles/04_erik_naiman_herst.pdf (last accessed May 31, 2017).
11. Piskaryov, D. (2017), *Calculating the Hurst exponent*, available at: https://www.mql5.com/ru/articles/2930 (last accessed May 31, 2017).

Надійшла (received) 19.06.2017
Прийнята до друку (accepted for publication) 18.10.2017

Дослідження методів вторгнення в комп'ютерні системи, засноване на показнику Херста

С. Ю. Гавриленко, В. В. Челак, М. Білогорський

Предметом дослідження даної статті є аналіз методів виявлення вторгнень в комп'ютерні системи. **Мета** статті – розробка ефективних методів і технологій протидії комп'ютерним вірусам. **Завдання**: дослідження сучасних засобів антивірусного захисту комп'ютерних систем; дослідження показника Херста для оцінки стану комп'ютерної системи; розробка програмної моделі оцінки стану комп'ютерної системи, що базується на показнику Херста, аналіз отриманих експериментальних даних. Використовуваними **методами** є: оцінка самоподібності процесу на основі показника Херста. **Отримані такі результати**. В роботі запропоновано метод фіксації аномального поведінки комп'ютерної системи на підставі показника Херста, який базується на аналізі завантаження центрального процесора і оперативної пам'яті комп'ютера. Результати досліджень показали, що вплив ряду вірусів на комп'ютерну систему призводить до прагнення показника Херста до середнього значення 0,5, що вказує на випадковість процесу. **Висновки.** Експериментальні дослідження підтверджують можливість використання показника Херста, як складової частини системи виявлення вторгнень в комп'ютерні системи.

Ключові слова: комп'ютерна система; несанкціонований доступ; комп'ютерний вірус; шкідливе програмне забезпечення; фрактальний аналіз; показник Херста.

Исследование методов вторжения в компьютерные системы, основанное на показателе Херста

С. Ю. Гавриленко, В. В. Челак, М. Белогорский

Предметом исследования в данной статье является анализ методов обнаружения вторжений в компьютерные системы. **Цель статьи** – разработка эффективных методов и технологий противодействия компьютерным вирусам. **Задачи:** исследование современных средств антивирусной защиты компьютерных систем; исследование показателя Херста для оценки состояния компьютерной системы; разработка программной модели оценки состояния компьютерной системы, базирующейся на показателе Херста, анализ полученных экспериментальных данных. Используемыми **методами** являются: оценка самоподобия процесса на основе показателя Херста. Получены **следующие результаты.** В работе предложен метод фиксации аномального поведения компьютерной системы на основании показателя Херста, который базируется на анализе загрузки центрального процессора и оперативной памяти компьютера. Результаты исследований показали, что влияние ряда вирусов на компьютерную систему приводит к стремлению показателя Херста к среднему значению 0,5, что указывает на случайность процесса. **Выводы.** Экспериментальные исследования подтверждают возможность использования показателя Херста, как составной части системы обнаружения вторжений в компьютерные системы.

Ключевые слова: компьютерная система; несанкционированный доступ; компьютерный вирус; вредоносное программное обеспечение; фрактальный анализ; показатель Херста.