Research of Phishing-Attacks Dynamics by the Wavelet-Analysis Method

Zhanna Deineko
Department of Media Systems and Technologies
Kharkiv National University
of Radio Electronics
Kharkiv, Ukraine
zhanna.deineko@nure.ua

Diana Draz
Student of the Department of Media Systems and Technologies
Kharkiv National University
of Radio Electronics
Kharkiv, Ukraine
diana.draz@nure.ua

Дослідження динаміки Phishing-атак методом вейвлет-анализу

Жанна Дейнеко
кафедра медіасистем і технологій
Харківський національний університет
радіоелектроніки
Харків, Україна
zhanna.deineko@nure.ua

Діана Драз
студент кафедри медіасистем і технологій
Харківський національний університет
радіоелектроніки
Харків, Україна
diana.draz@nure.ua

***Abstract*—In the article, the problems of research and forecasting of Phishing-attacks by a method of discrete wavelet-transformation are considered. A discrete wavelet transformation method for estimating of the probability of Phishing-attacks is proposed; features of using the basic characteristics of wavelet-analysis for the study of time series are shown. On the example of real data, confirmed by Phishing-attacks, a long-term dependence is revealed, which can be used to justify the forecasts of the occurrence of Phishing attacks.**

***Анотація*—У статті розглянуті питання дослідження і прогнозування Phishing-атак методом дискретного вейвлет-перетворення. Пропонується метод дискретного вейвлет-перетворення для оцінки ймовірності Phishing-атак, показані особливості використання основних характеристик вейвлет-аналізу для дослідження часових рядів. На прикладі реальних даних, підтверджених Phishing-атак, виявлена довгострокова залежність, що може бути використано для обґрунтування прогнозів виникнення Phishing-атак.**

Keywords—wavelet-analysis; wavelet-energy; time series, Phishing-attack

Ключові слова—вейвлет-аналіз; вейвлет-енергія; временной ряд; Phishing-атаки

I. INTRODUCTION

Recently, in the information field of modern remote communication systems, it is necessary to take into account the security of transmitting and receiving various data that may be accompanied by the presence of malicious components in the transmitted content [1]. Moreover, all sorts of technological tricks, along with objective shortcomings of software developers and designers, leave a large number of loopholes that somehow need to be taken into account. Among these loopholes a special place is occupied by the notion of phishing [2].

Today the most important risk and challenge is online fraud and phishing attacks. Phishing attacks were always used by attackers to steal users' passwords and electronic codes in virtual environment.

The term Phishing -attack is far from new and represents a process of deceiving customers for the subsequent theft of their identity and the transfer of their confidential information for criminal use. Access to logins and passwords is carried out by sending emails on behalf of popular brands and personal messages [2, 3].

Phishing attacks are categorized into different groups based on their attack method which include deceptive phishing, malware-based phishing, phishing by hosts file poisoning, system reconfiguration attacks, DNS-based phishing, content-injection phishing, search engine phishing, domain hijacking, PDF file phishing, and spear phishing. In all these phishing attacks, attackers use fake websites similar to main ones to achieve their goals and steal victims' information. In order to confront phishing attacks, different techniques have been invented which are called anti-phishing. Each anti-phishing

technique is focused on a certain characteristic of phishing pages or fake emails and some of them consider two or more characteristics and identify phishing pages using these characteristics.

II. Study of Phishing-attack by the Discrete Wavelet-Transformation method

The article considers the possibility and shows the effectiveness of conducting a study of Phishing-attack using the methodology of discrete wavelet analysis.

In terms of dynamics of Phishing-attacks, there is a data that characterize the daily number of Phishing-attacks. Such a data allows you to consider the power of Phishing-attacks in their daily recalculation over a period of time, the intensity of such attacks at certain intervals, or their consistency. At the same time, such conclusions can be useful both for their prevention and for predicting the possible growth of Phishing-attacks.

The processing and analysis of the sequence of data presented in the form of a time series is one of the most common methodologies in the study of various processes and phenomena related to different fields of activity and research. In this case, many time series generated by information flows have fractal properties and can be considered as stochastic fractals. Identification and study of such fractal properties can be carried out on the basis of the methodology of wavelet analysis [1].

In order to study the dynamics of Phishing attacks, the data from the website http://www.phishtank.com was taken, where data on confirmed Phishing attacks (Figure 1) and data on the total number of attacks were taken as the initial data. These data cover the period of time from October 1, 2012 to May 31, 2018 in their daily submission. On Fig. 1 and Fig. 2, along the abscissa, the days of the explored period are plotted, and along the ordinate axis, the number of confirmed and submitted Phishing-attacks, respectively. In terms of the study, graphs of autocorrelation functions of time series were presented. In both cases, a process of slowly decreasing autocorrelation function is observed. Consequently, the considered dynamics of the number of Phishing attacks is a self-similar process [1].

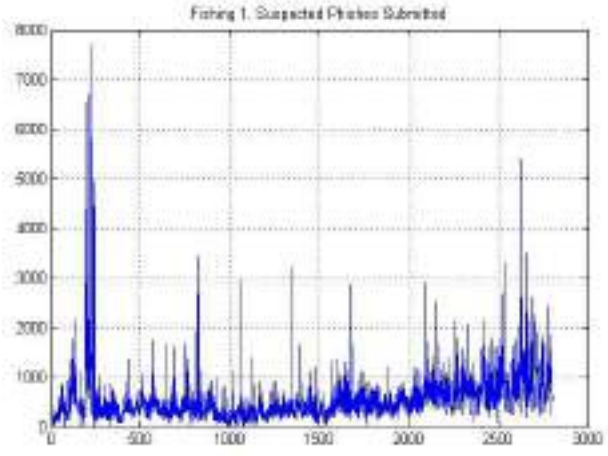

Fig. 1. Confirmed Phishing attacks data

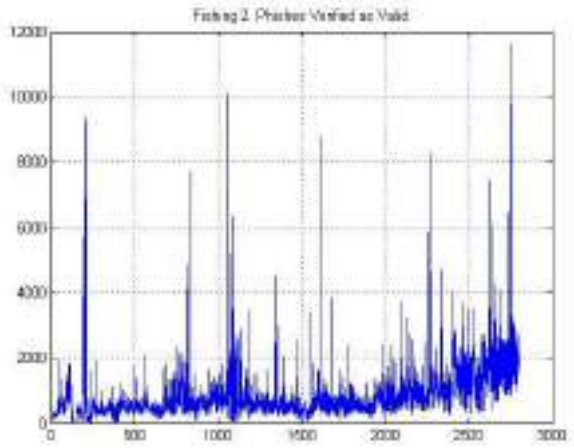

Fig. 2. Data on the total number of Phishing attacks

In the paper, general provisions concerning the wavelet analysis of time series were considered, and the features of using the basic characteristics of wavelet analysis for the study of time series were shown. Based on the real data, which are characterized by the time series presented, the ideology of applying wavelet analysis in the study of Phishing attacks is considered.

In particular, it is shown that the explored data series are characterized by the presence of a significant long-term dependence. This indicates the presence of long-term memory in the explored series of data, which makes it possible to build adequate forecast models for such a data.

III. Conclusion

Analysis of the spectrum of wavelet energy allows us to consider the presence of a trend component in the structure of the explored series of data, which can also be used to justify the forecasts regarding the occurrence of Phishing-attacks. At the same time, a mutual study of the spectrum of the wavelet energy of data series allows us to draw conclusions about the conformity of the methods used to identify the diversity of various events taking place in the Internet environment from the point of view of identifying potential Phishing attacks.

References

[1] Deineko, Zh. Properties of wavelet coefficients of self-similar time series // V. Lyashenko, Zh. Deineko, M. Ahmad / International Journal of Scientific and Engineering Research. – 2015. – 6 (1). – P. 1492-1499.

[2] Dadkhah, M. Prediction of phishing websites using classification algorithms based on weight of web pages characteristics // M. Dadkhah, M. Jazi, V. Lyashenko / Journal of Mathematics and Technology. – 2014. – 5(2). – P. 24-35.

[3] Dadkhah, M. Methodology of the Chaos Theory in Research of Phishing Attacks // M. Dadkhah, M. Jazi, V. Lyashenko / International Journal of Academic Research. – 2015. – 7(1). – P. 169-175.