EasyChair Preprint

№ 4580

Development of Algorithm for Improving Accuracy of Probability Coefficient of Threat Implementation in Personal Data Information Systems

Sergey Verevkin, Ksenia Naumova, Tatiana Tatarnikova, Pavel Bogdanov and Ekaterina Kraeva

November 16, 2020

DEVELOPMENT OF ALGORITHM FOR IMPROVING ACCURACY OF PROBABILITY COEFFICIENT OF THREAT IMPLEMENTATION IN PERSONAL DATA INFORMATION SYSTEMS

Verevkin S.A.[1], Naumova K.S.[1], Tatarnikova T.M.[0000-0002-6419-0072], Bogdanov P.Y.[0000-0002-7533-7316], and Kraeva E.V.[1]

[1] Russian State Hydrometeorological University, 79, Voronezhskaya st., 192007 St. Petersburg, Russia
vrjovkin@rambler.ru
ksenia.naumovaks@gmail.com
tm-tatarn@yandex.ru
45bogdanov@gmail.com
kate.smitt.by@mail.ru

Abstract. The article discusses the process of creating an algorithm that allows you to put the existing methodology for determining current threats to the security of personal data in detail when processing them in personal data information systems. The algorithm is based on the collection of intelligence from public sources (OSINT) for the purpose of its further analysis and use to determine the appropriate coefficients of the current methodology and for the further building of the information protection system.

Keywords: OSINT, corporate networks, security analysis, information security.

1 Introduction

Today, a matter of necessity of the need to ensure the information security of the organization is increasingly arose not only by large corporations and government entities, but also by small private organizations. The main reason for it is the increase in the cost of processed information in the networks of organizations that has become the most desirable resource of cybercriminals.

With the need to protect the information being processed, it is necessary to properly assess the current state of security of the information system in accordance with the requirements of current federal laws and other governing documents of supervisory bodies.

2 Justification of the existing problem

In carrying out the task of building an information protection system in organizations closely related to the processing of client databases that include personal data, an important point is the

need to determine the current personal data threats when processing them in ISPD in accordance with the current FSTEC methodology[1].

As a result of the actions described in this methodology, employees of the organization are faced with the task of determining numerical coefficients Y_1 and Y_2, which indicate the state of the initial security and the probability of the threat implementation.

Unlike the first coefficient determined by the table in the methodology, the value of the Y_2 coefficient should be determined by using the proposed verbal estimates corresponding to small, medium, high and unlikely.

It is worth noting the difficulty of conducting such assessments in the absence of any actual data on the current state of the organization's information systems and not to mention a further similar process for assessing the feasibility of a threat, which requires an impartial assessment of the possibility of implementing security incidents, including by the organization's staff.

3 Algorithm development

As a way to solve the problem of correctly determining the values of the Y_2, coefficient, we will build an algorithm that allows using open source software used for intelligence based on open information sources (OSINT) to search for existing threats.

Among the methods for conducting OSINT, the four-stage cyclic method for conducting data collection has gained the greatest popularity:

1) Definition of information search criteria
2) Retrieving searched data from open sources
3) Analysis of the received information
4) Structuring the obtained information in order to use it for further data search.

Therefore, the accuracy of the research conducted depends on the number of OSINT cycles, which allows you to determine the depth of analysis of the collected information depending on its type, secrecy and the wishes of the organization's management [2].

An important feature of OSINT is the full analysis of the organization's information and personnel resources. For this reason, we highlight three main steps of the algorithm being developed and consider the most successful methods of their implementation:

1) Analysis of public pages of the organization

It includes the collection and analysis of information about the organization posted in such sources of information, advertisements, organization websites, resources, tax information and other sources of information that allow you to obtain initial data on the activities of the organization: organizational structure, position, etc.

There are many software solutions, but as an example, we will consider the Maltego, which provides a convenient interface for visualizing data found and connections between it. Despite the fact that Maltego has a free version, the most effective are paid versions of the program that allow expanding its capabilities by connecting additional third-party libraries, the work of which is implemented by connecting using API keys. An example of analysis and construction of connections of collected data of the Russian State Hydrometeorological University (RSHU) website (rshu.ru) is shown in Figure 1.

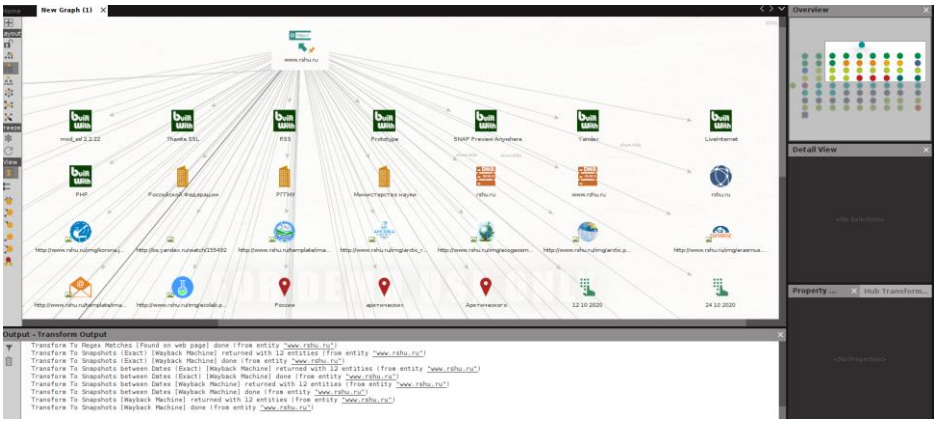

Fig. 1. Result of data collection from RSHU website (rshu.ru)

As a result of the analysis, it becomes possible to obtain the following information: contact information of the owners of network resources, hosting on the basis of which the organization's website is located, personal data of employees whose numbers are indicated on the website, information about the current and completed judicial proceedings of the organization and information about the dates of important events, such as: company management's birthdays, dates of corporate events and many other information that will further facilitate the receipt of additional information[3].

2) Analysis of employee information

In this step, you search for existing employees in your organization using the data you have received in the previous step. The main goal is to collect information about the largest number of employees in the organization using the previously obtained data. As a result of the analysis, it becomes possible to determine most of the employees of the organization with high accuracy through the analysis of social networks of these employees, their personal e-mails, phone numbers, home addresses and relationships between the employees.

We will use the OSINT Framework, which combines a huge number of solutions in the field of searching for information from open sources. The Maltego that was discussed earlier can also be used for these purposes, but most of its functionality for analyzing social networks used in Russia requires purchase of paid packages. The main advantage of the OSINT Framework is the ability to get the user the access to the maximum number of information from free sources, with additional indication of paid resources. Figure 2 shows the OSINT Framework options for Social Network and Mail Address Analysis.

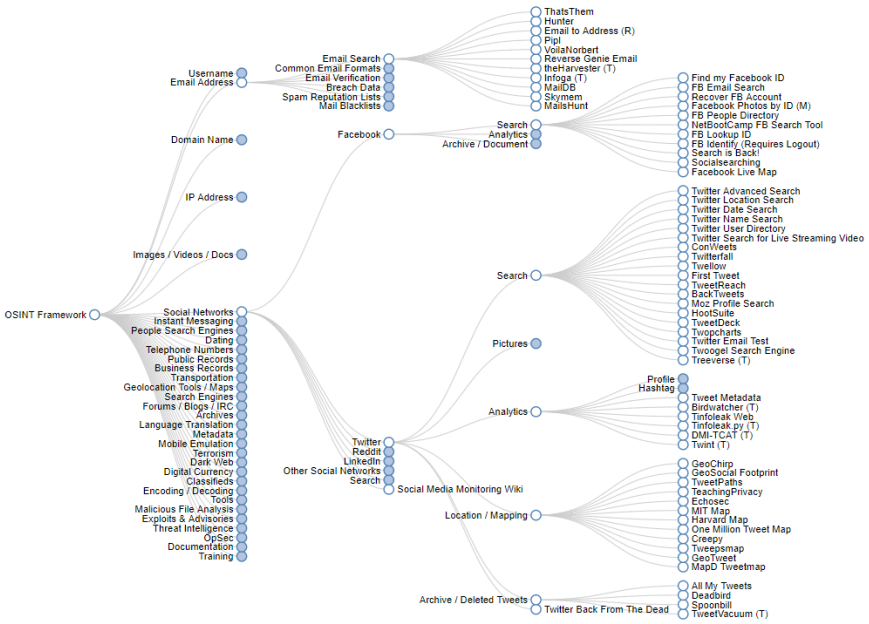

Fig. 2. OSINT Framework solutions for finding information on popular social networks and e-mail services

An important task of this step is to identify dissatisfied employees who openly express dissatisfaction with colleagues and the organization as a whole. Often, it is a dissatisfied employee who is a potential victim of social engineers who provoke the employee to help achieve their own goals.

3) Analysis of the organization's network

The last but no least important step is to analyze the current state of security of corporate networks of the organization. In this step, it is important to analyze the network infrastructure used by the system and application software, the security tools used, protocols and other information that allows the abuser to plan attacks for specific network components.

The task of analyzing data about an organization's network can be solved in many different ways, the application of which depends on the type of network and the devices used in it. One of the most famous tools is Nmap. Using Nmap to the address found using Maltego IP, we can get information about the system software used, which is used on the hosting network resource. Figure 3 shows the result of the website rshu.ru hosting operating system definition.

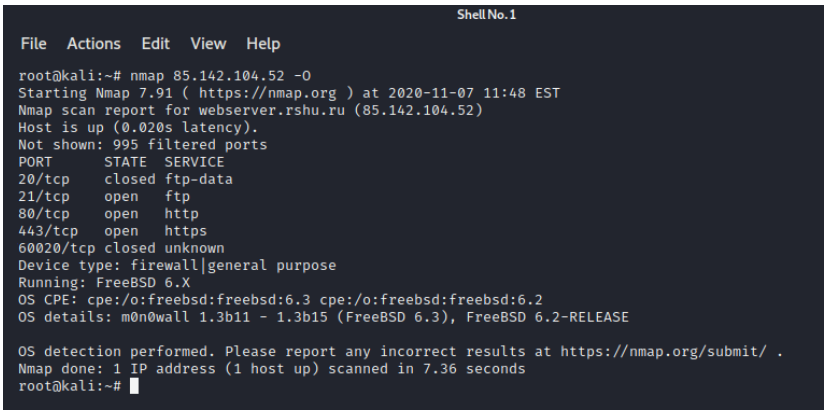

Fig. 3. Definition of the website rshu.ru hosting operating system

The main criterion for choosing an implementation tool is to locate an attacker in relation to the network of the organization. If located in a segment of the corporate network, the use of sniffers to analyze network traffic for the use of vulnerable network protocols is needed. At the same time, for the purpose of further penetration, it is necessary to use vulnerability scanners and Nmap analogues to search for vulnerabilities of border nodes of the network or to obtain information about the protection used in case of remote scanning of devices at the border of the investigated network in case of firewalls[4].

4 Conclusion

The Final report was created, as a result of each of the above described steps of the developed algorithm. This report is also a result of an in-depth detailing (number of cycles) of the conducted OSINT, which is a consequent of an analysis of the collection of collected data, on the basis of which, in the future, it becomes possible to further assess the probability of threat implementation using the established method of verbal gradations.

References

1."Methodology for determining current threats to personal data security during their processing in personal data information systems" FSTEC of 14.02 2008
2.Penetration Testing Execution Standard (PTES), URL: http://www.pentest-standard.org/index.php/Main_Page
3.«Maltego Desktop Application Guide»URL: https://docs.maltego.com/support/solutions/articles/15000008703-client-requirements#network-requirements-0-3
4.Tatarnikova T.M., Volskiy A.V. Estimation of probabilistic-temporal characteristics of network nodes with traffic differentiation//Informatsionno-Upravliaiushchie Sistemy. 2018. V. 94 No. 3. P. 54-60. DOI 10.15217/issn1684-8853.2018.3.5
5.Tatarnikova T.M. Statistical methods for studying network traffic //Informatsionno-Upravliaiushchie Sistemy. 2018. V.96. No.5. P. 35-43. DOI: 10.31799/1684-8853-2018-5-35-43