Botnet Detection Approach Based on DNS

Sergii Lysenko[a], Kira Bobrovnikova[a], Bohdan Savenko[a], Piotr Gaj[b], and Oleg Savenko [a]

[a] *Khmelnitsky National University, Khmelnitsky, Ukraine*
[b] *Silesian University of Technology, Gliwice, Poland*

Abstract

Botnets that use DNS technology are a serious threat on the Internet today. The potential of botnets is very large, from the spread of malware, ransomware, spam mailings to the theft of confidential information and money from bank accounts. Analysis of known methods and means of identification of botnets that use DNS has demonstrated the insufficient level of detection capacity of this type of botnets. Therefore, it is necessary to improve the method of botnets detection. The paper presents botnet detection approach based on DNS. The paper proposes a method of identifying botnets that use DNS based on the Decision Tree classifier with the application of the AdaBoost algorithm. The method allows you to ensure the detection of botnets that use DNS based on the characteristics of this technology of malicious software. The Decision Tree application algorithm is argued by the fact that it is a powerful tool for classification and prediction, and to strengthen the work of the above classifier, the AdaBoost algorithm was used in the study.

Keywords 1

DNS, Botnet, Cyberattack, Computer Network, Cybersecurity, Computer system, Malware, Malicious traffic, Botnet Detection, AdaBoost, Decision Tree

1. Introduction

The problem of information protection today is relevant, since it has no final solution and due to the rapid development of technology, new types of threats are constantly emerging. Modern computer systems typically rely on a domain name system (DNS). However, cybercriminals often abuse domain names because DNS traffic is usually unfiltered or allowed through a firewall, thereby providing a stable and seamless communication channel [1].

The relevance of the work lies in the development of an approach of identifying botnets that use DNS. The detection of new previously unknown threats should be based on a combination of all knowledge about botnets that use DNS-based evasion technologies. The use of this evasion technology can be detected by analyzing the signs removed from DNS messages using machine learning [2, 3]. The method should ensure the detection of botnet attacks that use DNS in the early stages or even before their occurrence [4, 5].

The purpose of scientific work is to increase the reliability of the process of identification of botnets using the technology of "domain flux", based on the analysis of DNS traffic.

The goal is achieved by solving the following main tasks:

1. To investigate the peculiarities of the functioning of botnets using the technology of "domain flux", taking into account the domain name system.
2. Analyze modern methods and means of identifying botnets based on DNS traffic in order to determine ways to increase the reliability of botnet detection.

IntelITSIS'2022: 3rd International Workshop on Intelligent Information Technologies and Systems of Information Security, March 23–25, 2022, Khmelnytskyi, Ukraine

EMAIL: sirogyk@ukr.net (S. Lysenko); bobrovnikova.kira@gmail.com (K. Bobrovnikova); savenko_bohdan@ukr.net (B. Savenko); piotr.gaj@polsl.pl (P. Gaj); savenko_oleg_st@ukr.net (O. Savenko)
ORCID: 0000-0001-7243-8747 (S. Lysenko); 0000-0002-1046-893X (K. Bobrovnikova); 0000-0001-5647-9979 (B. Savenko); 0000-0002-2291-7341 (P. Gaj); 0000-0002-4104-745X (O. Savenko).

3. Develop an appropriate model of the botnet, taking into account the domain name system, the DNS traffic model and the model of the process of identification of the "domain flux" evasion technology based on the analysis of DNS traffic.

4. Develop an improved method of identifying botnets that use "domain flux" technology by filtering DNS traffic using an accumulated database of whitelists, using the method of frequency lexical analysis of domain names, collecting signs of DNS traffic and analyzing them based on the Decision Tree classifier with the application of the AdaBoost algorithm.

Develop a botnet identification system that will improve the reliability and efficiency of the process of detecting botnets that use the "domain flux" technology.

The object of research is the process of identifying botnets that use the "domain flux" technology, based on the analysis of DNS traffic.

The subject of research is models, methods and software tools for identifying botnets using "domain flux" technology, based on the analysis of DNS traffic.

The aim of the work is to increase the reliability of the process of identification of botnets using the technology of "domain flux".

Scientific novelty of the obtained results. As a result of the study, the method of identification of botnets using the "domain flux" technology has been improved, based on the analysis of DNS traffic, which, unlike the well-known ones, uses complex analysis of DNS traffic using the Decision Tree classifier and the AdaBoost algorithm.

The practical significance of the results is the developed system of identification of botnets using the technology of "domain flux", based on a comprehensive analysis of DNS traffic with high reliability and efficiency based on the use of the RapidMiner platform, which is able to detect botnets with high reliability.

2. Related works

2.1. Botnets and DNS

Botnets play an important role in the spread of malware, and they are widely used to spread malicious activity on the Internet. The study of the literature shows that a large subgroup of botnets uses DNS to spread malicious actions and that there are different methods for detecting them using DNS queries. The Domain Name System (DNS) is a system that establishes a correspondence between the IP address and the domain name (and vice versa), and is designed to respond to DNS queries using the corresponding protocol [6].

Botnet is a computer network that includes a finite number of hosts that works with standalone software – running bots. Typically, a bot as part of a botnet is a program that is installed hidden on the victim's computer and allows the attacker to access the resources of the infected computer [7] and perform certain actions. Most often used for illegal activities – sending spam emails, collecting passwords on a remote system, organizing denial of service attacks, having access to personal information about users, theft of credit card numbers and access passwords [8]. "Domain flux" is a method of ensuring malicious activity of the botnet by constantly changing the domain name of the C&C server. Domain names are replaced in time based on the application of an algorithm that is known only to an attacker (botmaster). This makes it impossible to detect malicious traffic generated by the botnet.

Modern botnets such as Zeus (Zbot, PRG, Wsnpoem, Gorhax, Kneber, Chthonic, Panda), Torpig, Kraken, Conficker (DownUp, DownAndUp, DownAdUp, Kido), Mirai, "Star Wars" Twitter, Satori IoT, Trickbot, Emotet, usually use technology called "domain flux" and domain generation algorithm (DGA) [9] to generate a large number of pseudo-random domain names to dynamically manage network operator bots and their bots.

Typically, botnets generate a large number of DNS queries registered to the same IP address, and they often generate many failures in DNS traffic. Failed DNS queries may indicate the presence of bots on clients, while successful queries that occur in time next to unsuccessful ones are likely related to benign user.

In addition, the botnet can quickly transmit messages to all bots, which is one of the main advantages. DNS traffic monitoring is an important task and helps to detect botnets that use the DNS analysis [10].

Identifying domain names generated algorithmically through DNS traffic analysis has different advantages.

For example, DNS includes only a small amount of traffic throughout the network, making it appropriate for analysis even on large large-scale networks. Additionally, DNS traffic is typically cached, which reduces network load. Moreover, the analysis of DNS queries helps to detect attacks in the early stages or even before they occur.

2.2. Related works

Today, the scientific community has developed a large number of methods for solving the problem of identification of botnets using the technology of "domain flux".

In particular, [11] describes a method that is based on the analysis of DNS queries, determines the correlation of various logs and error messages, diagnoses based on the history of suspicious activity, detects the activities of botnets based on DNS traffic failure and diagnoses based on DNS group activity, and analyzes group activity on the network.

In [12] a method is proposed to identify DGA based on the dictionary using graph theory.

The article [13] describes the method of detecting bot modes using frequency analysis of character distribution and weighted domain name scores.

Paper [14] presents a method for detecting botnets by classifying text strings of domain names based on n -grams.

The paper [15] provides a method for detecting botnets based on the analysis of DNS traffic functions. This method passively captures all DNS traffic from the gateway network, and then extracts key functions to identify pseudo-random domain names.

The anomaly detection and passive DNS analysis approach for botnet detection are presented in [16, 17]. In [18] Identifying legitimate Web users and bots with different traffic profiles - an Information Bottleneck approach is presented.

Thus, in contrast to heuristic methods, machine learning-based methods achieve high accuracy in detecting botnets using "domain flux" technology. Of course, if in the process of generating pseudo-random domain names, the logic of work changes, as well as certain characteristics change dramatically, which may significantly reduce the efficiency of detection of botnets of this type.

The anti-virus tools in question detect all known types of malware using signature databases and existing heuristic approaches. However, none of the popular anti-virus tools today detect 100% of botnets that use DNS.

2.3. Conclusions and the problem statement

Known methods of botnets identification that use DNS are the subject to a decrease in the accuracy of detection of new unknown botnets of this type when using pseudo-random domain names other than known methods of generating pseudo-random domain names. Also, most methods do not allow detecting attacks in the early stages or even before they occur.

Therefore, in order to increase the reliability and efficiency of the process of identifying botnets that use DNS, it is necessary to develop or improve the method based on a comprehensive analysis of DNS traffic through the use of machine learning algorithms.

3. Model of the process of identification of botnets that use DNS

3.1. Process Formalization for the functioning of botnets that use DNS

Botnets are distributed over the Internet using similar approaches to other SDPs. They can spread like a worm, or they can disguise themselves as a Trojan executing following operations: the botmaster issues a command (set of parameters and settings) to the command-and-control center to carry out the attack; in turn, the command-and-control center sends a message to all bots of the botnet, which immediately begin to execute the commands of the botmaster.

Consider the model of a botnet that uses DNS as a control system for infected bots of a computer system and present it in the form of a tuple:

$$M_{DF} = \langle C, A, P, B, L, F \rangle, \tag{1}$$

where C – a set of command and control servers (C&C) of the botnet;

N_{CC} – the number of command-controlling servers of the botnet, $C = \{c_j\}_{j=1}^{N_{CC}} N_{CC}$;

$A = \{a_j\}_{j=1}^{3}$ – type of botnet architecture;

$P = \left\{p_j^{port}\right\}_{j=1}^{N_P}$ – a set of network protocols used for the functioning of the botnet – the number of network protocols used by the botnet,

where N_P– a set of ports used for commissioning with the botnet, where $N_P port \in NPortNPort = \{1..65535\}$27];

$B = \{b_j\}_{j=1}^{N_B}$ – a set of bots of the bot-network,N_B – the number of bots included in the bot network;

$L = \{l_j\}_{j=1}^{5}$ – a set of stages of the life cycle of the botnet;

$F = \{f_j\}_{j=1}^{N_F}$ – a set of bot functions that can be performed in the corresponding phase of the botnet life cycle,

N_F – the number of functions that bots can perform botnets.

3.2. Model of attack carried out by a botnet that use DNS

Consider the botnet attack model, which uses DNS as a set of commands to carry out malicious activities that can be performed by bots of botnets and their possible use scenarios:A

$$A = \left\{a_{DDoS}, a_{spam}, a_{phishing}, a_{phishing}, a_{espionage}, a_{posting}, a_{proxy}\right\}, \tag{2}$$

where a_{DDoS} – distributed denial of service attack;

a_{spam} – spam attack;

$a_{phishing}$ – phishing attack;

$a_{espionage}$ – espionage;

$a_{posting}$ – placement of harmful content, such as the placement of content or advertising;

a_{proxy} – carrying out attacks using proxy servers.

Botnets are distributed over the Internet using similar approaches to other malware. They can spread like a worm, or they can disguise themselves as a Trojan. Let's take a closer look at the model of ways to spread malware [24] and present it as follows:

$$a_{distribution} = \{d_{os}, d_s, d_a, d_{se}\}, \tag{3}$$

where d_{os} – distribution due to vulnerabilities in the operating system;

d_s – distribution through services and services;

d_a – distribution through applications and applications;

d_{se} – Spread through social engineering.

3.3. Botnet detection model

Since the detection of botnets that use DNS is based on the botnet model, taking into account the domain name system and DNS traffic model, it is an important task to develop a DNS traffic model and a DNS package.

Let's present a model of DNS traffic as a tuple:

$$DNS_{traffic} = \langle M, C, S, D \rangle, \tag{4}$$

where M – a set of DNS messages sent and received from a set of computer systems of the network, $M = M_O \cup M_I$, where M_O and M_I – are the set of outgoing and incoming DNS messages in the network, respectively;

C – a set of computer systems of the network;

S – a set of DNS servers to which DNS queries and DNS responses were sent and received, respectively $S = S_L \cup S_N$,

where S_L and S_N – set of local and non-local DNS servers, respectively;

D – a set of requested domain names by a set of hosts of the network,

where $D = \{d_j\}_{j=1}^{N_D} N_D$– the number of different domain names.

Let us present a model of DNS messages. It has to involve such elements as information concerning the fields of incoming DNS messages. It enables the detection of the botnet that use DNS.   that can be used to identify botnets that use DNS,

Thus, DNS messages may be presented:

$$R = \langle R_{Mac}, R_{IP}, R_{Port}, R_T, \langle R_H, R_{Req}, R_{Ans}, R_{Ath}, R_{Add} \rangle \rangle, \tag{5}$$

where R_{Mac} is the host MAC address;

R_{IP} – host IP address of the SOURCE of the DNS package;

R_{Port} – the source port of the DNS package;

R_T – the time of receipt of the DNS package;

R_H — header DNS-message section;

R_{Req} – question DNS message request section;

R_{Ans} – answer DNS message section;

R_{Ath} –authority DNS message section;

R_{Add} – additional information DNS message section.

The process of extracting signs from incoming DNS messages for a specific domain name is presented as a function:

$$f_{extr}(D, M, R, V) \to I, \tag{6}$$

where V – set of signs that indicate the botnet presence.

The set of signs indicating the activity of the botnet, which use DNS, consists of the following elements:

$$V = \{N_{dom}, S_{bit}, T_{ttl}, L_{dom}, N_{num}, W_{dom}\}, \tag{7}$$

where $N_{dom}$is the number of domain names that share an IP address;

S_{bit} – a binary sign of the success of the DNS query (if S_{bit}= false – a failed DNS query, and if S_{bit}= true – a successful DNS query);

T_{ttl} – TTL-period;

L_{dom} – the length of the domain name;

N_{num} – number of digits in the domain name;

W_{dom} – a balanced estimate of the frequency lexical analysis of domain names, determined by the formula:

$$W_{dom} = \frac{\sum_{i=0}^{n} X_i}{n}, \tag{8}$$

Where n – the number of letters in the domain name; X_i – frequency of use i-th letter.

Let's present a model of the process of identification of botnets that use DNS as follows:

$$P = \langle M_{DF}, DNS_{traffic}, f_{extr}, f_{map}, f_{clas}, f_{mes} \rangle, \quad (9)$$

where M_{DF}– botnet model; $DNS_{traffic}$ – DNS traffic model;

f_{extr} – function of the algorithm for extracting signs from incoming DNS messages; $f_{map}(DNS_{traffic}, I) \rightarrow V$ – function of sampling signs from DNS traffic;

$f_{clas}(DNS_{traffic}, R, V) \rightarrow Res$ – the function of classification of DNS messages of DNS traffic in the network;

$f_{mes}(Res) \rightarrow Mes$ – function of notification of detection of bots of bots.

4. Botnet Cyberattacks Detection Approach Based on DNS

4.1. Method for Botnet Detection Based on Decision Tree Classifier

The paper proposes a method of identifying botnets that use DNS based on the Decision Tree classifier with the application of the AdaBoost algorithm. The method allows you to ensure the detection of botnets that use DNS based on the characteristics of this technology of malicious software.

The Decision Tree application algorithm is argued by the fact that it is a powerful tool for classification and prediction, and to strengthen the work of the above classifier, the AdaBoost algorithm was used in the study.

The method consists of the following stages: preparation, training of the system and direct detection of the activities of the botnet that use DNS.

The preparation stage includes the following steps:

1) analysis, modeling and identification of key features that will be used to identify botnets that use DNS;

2) collection of test data (network traffic) for training.

The training stage includes the following steps:

1) downloading test data (network traffic);

2) data conversion – in most cases, the available data is not suitable for use directly for teaching the machine learning model, the necessary data must be pre-processed;

3) frequency lexical analysis of domain names;

4) formation of a database of white lists of domain names;

5) model training –using the Decision classifier and the AdaBoost algorithm, based on the signs identified at the preparation stage for identifying botnets that use DNS;

6) evaluation of the model.

The stage of detection of the activities of botnets that use DNS includes the following steps:

1) monitoring of network traffic;

2) filtering DNS traffic that uses weeding out known DNS queries that contain legitimate domain names;

3) Collect all available parameters and features in filtered collected traffic

4) Identify groups in which the DNS query is unsuccessful.

5) identification of queries in which domain names by statistical analysis method are most likely formed algorithmically;

6) Comparing multiple groups of features and analyzing them using the Decision Tree classifier and the AdaBoost algorithm;

7) formation of conclusions.

To train the model, the Decision Tree classifier was used with the AdaBoost algorithm.

The pseudocode of the AdaBoost algorithm is given below:

```
We have: for all $(x_1, y_1),...,(x_m, y_m)$ $x_i \in X, y_i \in Y = \{-1, +1\}$.
Initialize $D_1(i) = \frac{1}{m}, i = 1,...,m.$
For every $t = 1,...,T$:
```

Find $h_t : X \rightarrow \{-1,+1\}$ a classifier that minimizes weighted classification error: $h_t = \arg\min_{h_j \in H} e_j$ where $e_j = \sum_{i=1}^{m} D_t(i)[y_i \neq h_j(x_i)]$.

If the value is , then $e_t \geq 0.5$ the stop is performed.

Select $\alpha_t \in R$, $\alpha_t = \frac{1}{2}\ln\frac{1-e_t}{e_t}$ where the weighted classifier failed. e_t h_t

$$D_{t+1}(i) = \frac{D_t(i)e^{-\alpha_t y_i h_t(x_i)}}{Z_t}, \tag{10}$$

where the normalization parameter is Z_t D_{t+1} $\sum_{i=1}^{m} D_{t+1}(i) = 1$(selected so that the probability distribution is selected, that is).

Build the resulting classifier:

$$H(x) = sign\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right). \tag{11}$$

The expression to update the distribution must be designed in such a way that the following condition is met:

$$e^{-\alpha_t y_i h_t(x_i)} \begin{cases} <1, & y(i) = h_t(x_i) \\ >1, & y(i) \neq h_t(x_i) \end{cases}. \tag{12}$$

Thus, after selecting the optimal classifier h_i for the distribution D_t of objects x_i that are a classifier identifies correctly, have weights less than those determined incorrectly.

In the initial step, DNS traffic is obtained by monitoring the network through the SPAN port of the network switch (Switched Port Analyzer), which duplicates packets from one or more ports to a separate port and for each domain name, the value of a weighted estimate of frequency lexical analysis is determined.

Weighted estimates of the frequency lexical analysis of DNS domain names are used to further identify botnets using "domain flux" technology, based on the W_{dom} Decision Tree classifier and the AdaBoost algorithm, as one of the signs indicating this application of evasion technology.

Then all selected and analyzed data from the filtered DNS traffic are combined into a set of features, which will allow detecting botnets that use DNS, based on the Decision Tree classifier with the application of the AdaBoost algorithm.

For all the collected grouped data, the applied conversion and normalization, resulting in a set of features for each domain name

$$V_j = \{N_{dom,j,}S_{bit,j}, T_{ttl,j}, L_{dom,j}, N_{num,j}, W_{dom,j}\},$$

where $j \in R'$– the number of collected DNS messages after filtering.

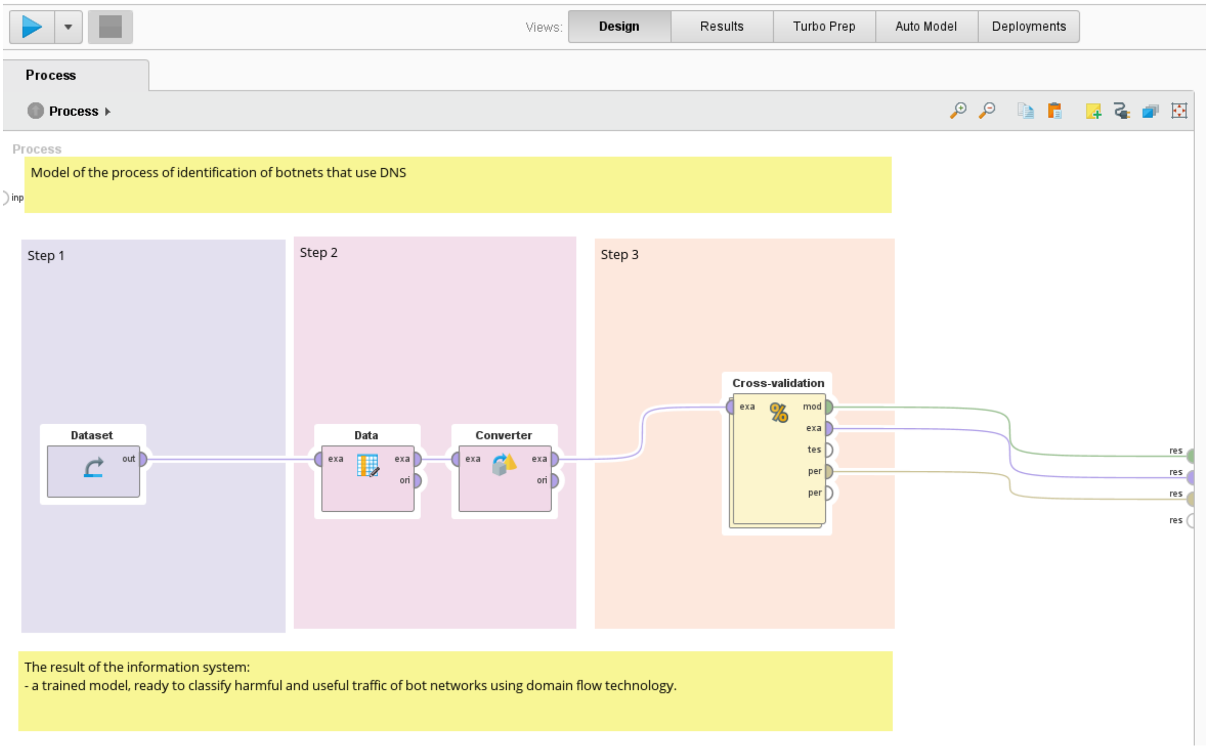
The last step is to form conclusions based on the analysis of a set of features V_j for each domain name by the AdaBoost machine learning algorithm. Since binary classification is used by a machine learning algorithm, the output obtains a result in accordance with each requested domain name, which can acquire two values: malicious request from the bot or benign request.

4.2. Implementation of the Botnet Detection system

In order to verify the effectiveness of the proposed method, a botnet identification system was implemented, which is based on the use of the RapidMiner open source platform [19].

RapidMiner is an integrated environment for data processing in large information arrays, machine learning, text analytics and construction of predictive models, as well as for solving applied and scientific problems.

Model of the botnet detection system using designed in a RapidMiner environment is presented in Figure 1.

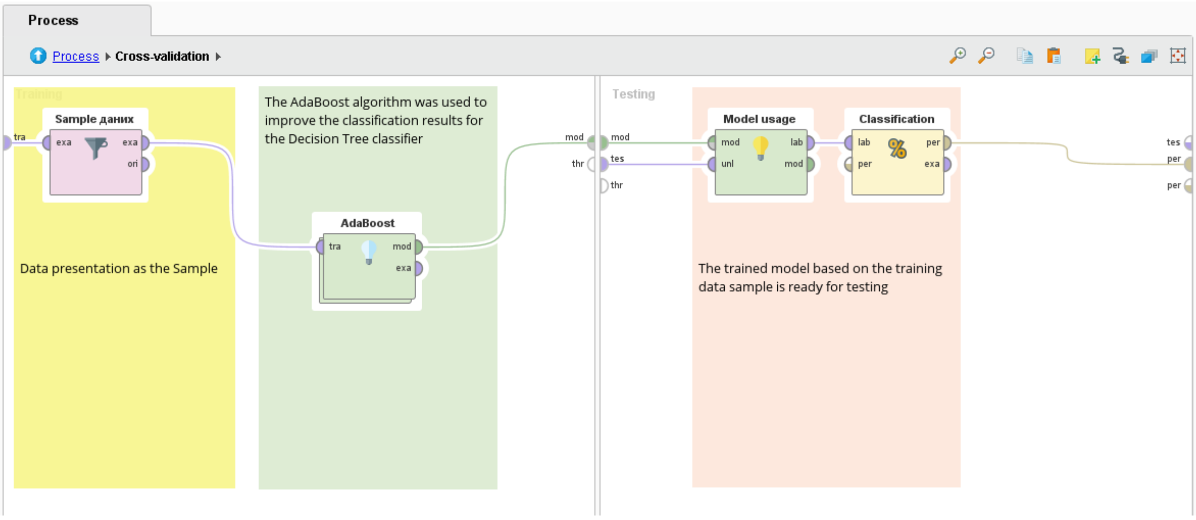


Model of the process of identification of botnets that use DNS

The result of the information system:
- a trained model, ready to classify harmful and useful traffic of bot networks using domain flow technology.

Figure 1: Model of the botnet detection system using designed in a RapidMiner environment

Subsystem of application of the AdaBoost algorithm is shown in Figure 2.



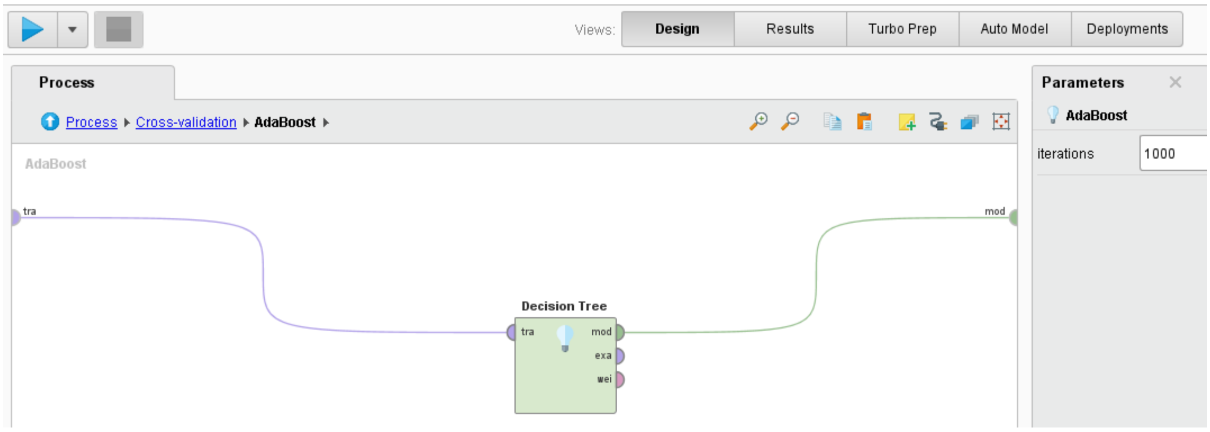
The AdaBoost algorithm was used to improve the classification results for the Decision Tree classifier

Data presentation as the Sample

The trained model based on the training data sample is ready for testing

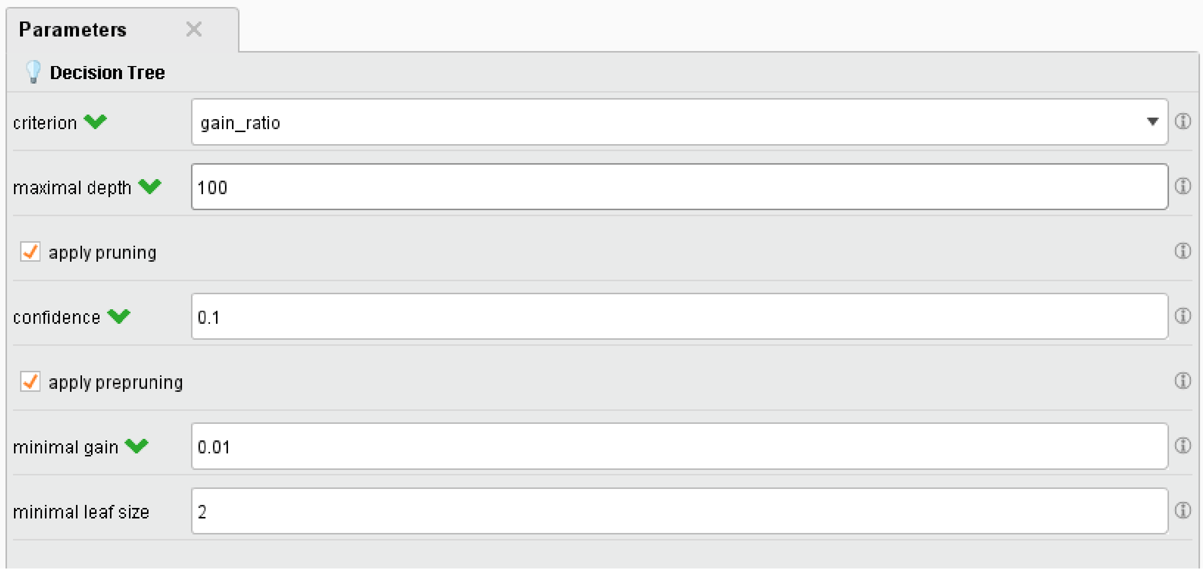
Figure 2: Subsystem of application of the AdaBoost algorithm

The process of strengthening the Decision Tree classifier by the AdaBoost algorithm is shown in Figure 3. Parameters for Decision Tree evaluation are presented in Figure 4.



iterations 1000

Figure 3: The process of strengthening the Decision Tree classifier by the AdaBoost algorithm


Parameters

Decision Tree

criterion ✓ gain_ratio

maximal depth ✓ 100

☑ apply pruning

confidence ✓ 0.1

☑ apply prepruning

minimal gain ✓ 0.01

minimal leaf size 2

Figure 4: Parameters for Decision Tree evaluation

4.3. Experimental studies of the effectiveness of the botnet detection system

To assess the effectiveness of proposed method of botnets detection, a number of experiments were carried out. the experimental environment is based on the framework described in [20-23].

To ensure unbiased results at the training stage, the dataset [24] was divided into two parts.

The first is 75% for training, and the remaining 25% is used to check the correctness of the system.

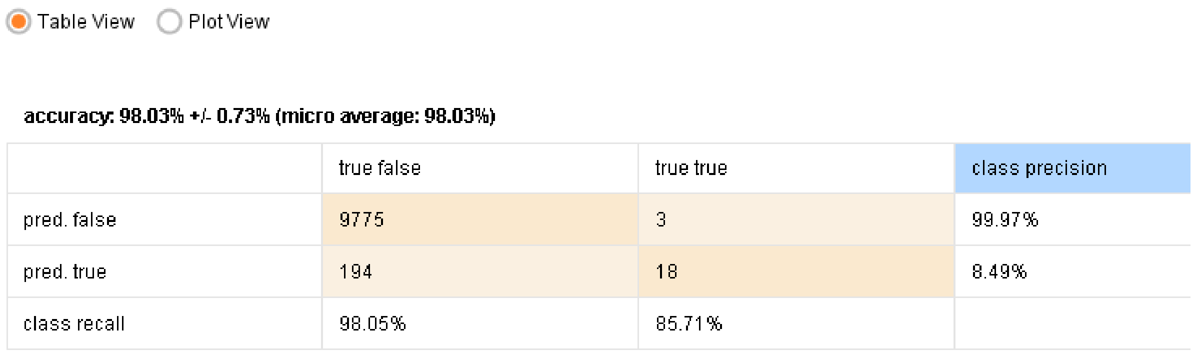
A total of 19,500 domain names (components of the training sample) were analyzed, among which 14,625 (75%) were selected for training, and 4,875 (25%) were used to verify correctness.

There were 9,969 requests for input experiments. There were also requests from bots of the bot network using the "domain flux" technology.

The total number was 306. Correctly identified 9775 requests, which is 98.05% of the total.

The total number of correctly and incorrectly identified requests is presented as a result of the system in Figure 5.

Thus, the proposed method demonstrated the possibility of identifying botnets using "domain flux" technology with high reliability (98.05%).



● Table View ○ Plot View

accuracy: 98.03% +/- 0.73% (micro average: 98.03%)

	true false	true true	class precision
pred. false	9775	3	99.97%
pred. true	194	18	8.49%
class recall	98.05%	85.71%	

Figure 5: Results of botnet detection system

5. Conclusions

The method of identification of botnets that use DNS has been developed. The results of the study obtained such scientific results.

The peculiarities of the functioning of botnets that use DNS taking into account the domain name system was investigated. Modern methods and means of identification of botnets based on DNS-traffic were analyzed in order to determine ways to increase the efficiency of botnet detection.

The corresponding model of the botnet was developed taking into account the domain name system, the DNS traffic model and the model of the process of botnet detection on the analysis of DNS traffic.

An improved method of identification of botnets that use DNS by filtering DNS traffic using the accumulated database of white lists, collecting signs of DNS traffic and analyzing them based on the Decision Tree classifier with the application of the AdaBoost algorithm has been developed.

A system for identifying botnets has been developed that will ensure an increase in the reliability and efficiency of the process of detecting botnets that use DNS.

Experimental research demonstrated the ability of the proposed method to identify botnets that use DNS with high reliability (up to 98.05%).

6. References

[1] Check point software cyber security report 2022. URL: https://www.ntsc.org (accessed on February 1, 2022).

[2] Wang Zishuo, Wang Chunyang, Ding Lianghua, Wang Zeng, Liang Shuning, Parameter identification of fractional-order time delay system based on Legendre wavelet, Mechanical Systems and Signal Processing, Volume 163, 2022, 108141, ISSN 0888-3270, https://doi.org/10.1016/j.ymssp.2021.108141.

[3] Zhang, Wang Huiqin, J Wang Chun, Meng Xudong Chen,. Integration of cuckoo search and fuzzy support vector machine for intelligent diagnosis of production process quality. Journal of Industrial & Management Optimization. 2017. 13. 10.3934/jimo.2020150.

[4] S.N.Thanh, M.Stcgc, P.I.El-Habr, J.Bang, N.Dragoni, Survey on botnets: Incentives, evolution, detection and current trends. Future Internet, 2021,13(8), 198.

[5] G.Suchacka, A.Cabri, S.Rovetta, F.Masulli, Efficient on-the-fly Web bot detection. Knowledge-Based Systems, 2021, 223,107074.

[6] RFC 1034, Domain Names - Concepts and Facilities.

[7] Lekssays, A., Landa, L, Carminati, B., Ferrari, E. PAutoBotCatcher: A blockchain-based privacy-preserving botnet detector for Internet of Things. Computer Networks. 2021, 200,108512.

[8] J.Shi., Y.-B.Leau, K. Li, J.H.Obit. A comprehensive review on hybrid network traffic prediction model. International Journal of Electrical and Computer Engineering, 2021, 11(2), pp. 1450-1459.

[9] D.Truong, G.Cheng Detecting domain-flux botnet based on DNS traffic features in managed network. Security Comm. Networks. 2016. P. 2338–2347. DOI: 10.1002/sec.1495.

[10] O.Olowoyo, P. Owolawi, Malware Classification using Deep Learning Technique. 2020 2nd International Multidisciplinary Information Technology and Engineering Conference, IMITEC 2020, 9334071.

[11] C.Nafari, E.Mahdipoor, Sayed Javadi H.Hajj Detection of active botnets based on DNS traffic analysis. Journal of Advances in Computer Engineering and Technology. 2019, Vol. 5, No. 3. P. 129–142.

[12] M.Pereira, S. Yu B.Coleman, M.D.Cock, A.C. Nascimento. Dictionary Extraction and Detection of Algorithmically Generated Domain Names in Passive DNS Traffic. 21st International Symposium : RAID 2018. Heraklion, September 10-12, 2018. DOI: 10.1007/978-3-030-00470-5_14.

[13] E.Agyepong, W.J.Buchanan, K.Jones. Detection of Algorithmically Generated Malicious Domain Using Frequency Analysis. International Journal of Computer Science and Information Technology. 2018. P. 91–111 DOI: 10.5121/ijcsit.2018.10306.

[14] T.Wang, L.C.Seidenberg. Detecting Algorithmically Generated Domains Using Data Visualization and N-Grams Methods. Proceedings of Student-Faculty Research Day, CSIS, Pace University, May 5, 2017.

[15] J.Spooren, D.Preuveneers, L.Desmet, P.Janssen, W.Joosen. Detection of algorithmically generated domain names used by botnets: a dual arms race. SAC '19: Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing. 2019. P. 1916–1923. DOI: https://doi.org/10.1145/3297280.3297467.

[16] H.Qin, J.Yang, X Luo. Z.Li, Q. Guo, Research on DNS anomaly detection technology based on multiple features. Journal of Shenzhen University Science and Engineering. 2020, 37. pp. 36-43.

[17] X.Guo, Z.Pan, Y.Chen, Application of Passive DNS in Cyber Security. Proceedings of2020 IEEE International Conference on Power, Intelligent Computing and Systems, CPICS 2020, pp. 2S7-259,9202344.

[18] G.Suchacka, J. Iwariski. Identifying legitimate Web users and bots with different traffic profiles - an Information Bottleneck approach. Knowledge-Based Systems, 2020,197,10587S

[19] RapidMiner’s data science platform. https://rapidminer.com/ (accessed on February 1, 2022).

[20] Tomas Sochor, Nadezda Chalupova. Interpersonal Internet Messaging Prospects in Industry 4.0 Era. In: Recent Advances in Soft Computing and Cybernetics. Springer, Cham, 2021. p. 285-295.

[21] O.Savenko, S.Lysenko, A.Kryschuk Multi-agent based approach of botnet detection in computer systems. Communications in Computer and Information Science. 2012. Vol. 291. PP.171-180, ISSN: 1865-0929.

[22] Oleg Savenko, Sergii Lysenko, Andrii Kryshchuk, Yuriu Klots. Botnet detection technique for corporate area network. Proceedings of the 7-th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, Berlin (Germany), September 12–14, 2013. Berlin, 2013. Pp. 363–368. ISBN 978-1-4799-1426-5.

[23] Sergii Lysenko, Kira Bobrovnikova, Serhii Matiukh, Ivan Hurman, Oleg Savenko. Detection of the botnets’ low-rate DDoS attacks based on self-similarity. International Journal of Electrical and Computer Engineering. 2020. Vol. 10., №4. PP.3651-3659, ISSN: 2088-8708.

[24] O. Savenko, A. Nicheporuk, I. Hurman, S. Lysenko. Dynamic signature-based malware detection technique based on API call tracing. CEUR-WS. 2019. Vol. 2393. P.633-643, ISSN: 1613-0073.

[25] Canadian Institute for Cybersecurity. Botnet dataset, https://www.unb.ca/cic/datasets/botnet.html (accessed 15.01.2022).

[26] IoT dataset. URL: https://github.com/thieu1995 /iot dataset (accessed on 15.01.2022).

[27] IoTPOT dataset. URL:https://ipsr.ynu.ac.jp/iot /index. html# datasets (accessed on 15.01.2022).