Detecting APT Attacks Based on Network Traffic Using Machine Learning

Cho Do Xuan

Information Assurance dept. FPT University, Hanoi, Vietnam
E-mail: chodx@fe.edu.vn

Received 30 September 2020; Accepted 31 October 2020;
Publication 17 February 2021

Abstract

Advanced Persistent Threat (APT) attacks are a form of malicious, intentionally and clearly targeted attack. By using many sophisticated and complicated methods and technologies to attack targets in order to obtain confidential and sensitive information. In fact, in order to detect APT attacks, detection systems often need to apply many parallel and series techniques in order to make the most of the advantages as well as minimize the disadvantages of each technique. Therefore, in this paper, we propose a method of detecting APT attacks based on abnormal behaviors of Network traffic using machine learning. Accordingly, in our research, the abnormal behavior of APT attacks in Network Traffic will be defined on both components: Domain and IP. Then, these behaviors are evaluated and classified based on the Random Forest classification algorithm to conclude about the behavior of APT attacks. Details of the definition of abnormal behaviors of the Domain and IP will be presented in Section 3.2 of the paper. The synchronous APT attack detection method proposed in this paper is a novel approach, which will help information security systems detect quickly and accurately signs of the APT attack campaign in the organization. The experimental results presented in Section 4 will demonstrate the effectiveness of our proposed method.

Journal of Web Engineering, Vol. 20_1, 171–190.
doi: 10.13052/jwe1540-9589.2019

Keywords: Advanced Persistent Threat, APT attack detection, Network traffic, domain, abnormal behaviour, machine learning.

1 Introduction

1.1 Introduction to APT Attack

APT attack technique is an advanced and targeted attack technique [1]. This is shown in its persistence and ability to conceal and hide [1, 2]. The studies [1–4] presented the definitions and concepts of terms: Advanced, Persistent, and Threat in this attack technique. Moreover, the studies [1, 2] pointed 4 characteristics that highlight the difference between APT attack and other network attack techniques including Targeted, Persistent, Evasive, and Complex. This difference plays a vital role in illustrating the APT attack much more dangerous than other attack techniques. The study [2] identified the phases of the APT attack campaign including Reconnaissance, Preparation, Targeting, Further access, Data gathering, and Maintenance. Currently, the APT attack is considered the most dangerous cyber-attack technique and causes many difficulties and damage to organizations and state agencies of countries around the world. Therefore, the problem of early detecting and warning this attack technique is very necessary today.

In this paper, we propose the APT attack detection method using a combination of two methods of analyzing signs and abnormal behavior of domain and IP in Network traffic. Accordingly, the characteristics of the APT attack detection method that we propose are as follows:

- Step 1: Extracting the behavior and features of APT attacks in Network traffic. At this step, we propose features and behavior of the domain and IP of APT attacks in Network traffic;
- Step 2: Detecting APT attack based on Network traffic using abnormal behavior analysis technique and the Random Forest (RF) machine learning algorithm. After having features that represent the difference between APT domains and APT IPs with clean domains and IPs, we use the Random Forest algorithm to evaluate and classify in order to detect suspect domains and IPs of APT attacks. Besides, in order to improve the efficiency of the APT IP detection process, we consider the output of the malicious domain detection process as a feature of the behavior of IP-based APT attacks. The combination of features extracted from the IP and the label of the domain makes the detection process more efficient due to the correlation between the domain and the IP.

1.2 Survey about APT Attack Detection Models

The studies [3–5] analyzed a number of difficulties and challenges that made APT attack detections were not highly efficient including the lack of public data on the APT attacks, the data imbalance, using standard coding protocols, etc. Besides, APT attacks are designed specifically for each specific target and object, so if rely on the experience and data of single attacks, it will not be able to detect new attacks. To overcome the above problems, the trend of APT attack detection systems is often designed to ensure maximum detection of signs and behaviors of APT attacks in both depth and width [1, 2]. Some recent popular models of detecting APT attacks include:

- Parallel model is a model using different algorithms to analyze and detect attacks from input data. All algorithms will report simultaneous detection results. After obtaining these detection results, correlation computation algorithms will be used to conclude if this is an APT attack or not [4, 5]. The advantage of the parallel model is that it provides more detection capabilities because each algorithm can focus on detecting the attack steps in an attack, and the decision system will ultimately decide whether an attack happens or not.
- Synchronous model is a complex model with a combination of series and parallel design. Other series or parallel elements in the system can analyze data from different sources for analyzing and detecting attacks. In this system, it is not required for each layer in the model to give a warning and the next layer to give warning higher than the previous model. Rather, this system creates dependence on relationships between events at different places and times instead of a single event. After obtaining these detection results, correlation computation algorithms will be used to conclude if this is an APT attack or not [4, 5].

1.3 Contributions of Paper

Our paper is presented as follows: the urgency of the research problem is presented in Section 1. In Section 2, we present the process of researching, surveying, and evaluating related works. The proposed model, the definition of abnormal features and behavior, and attack detection methods are presented in Section 3. Section 4 presents the results of the experimental process. Conclusion and evaluation are presented in Section 5. The practical significance and scientificity of our paper include:

- Proposing an APT attack detection model based on Network traffic using machine learning. Our proposed model uses a combination of both static and dynamic methods to optimize the time and efficiency of the monitoring and detection process.
- Proposing an APT attack detection method based on the technique of analyzing abnormal behavior of domains and IPs in Network traffic using the Random Forest machine learning algorithm. This is a novel proposal based on the correlation between Domain and IP in Network traffic. This combination technique will give better efficiency than individual detection techniques.

2 Related Works

2.1 Detecting APT Domain

In the paper [6], the authors detected APT attacks based on two factors: DNS log and Network traffic. For the APT attack detection technique based on the DNS log, the authors used 5 feature groups: Domain-based features, Time-based features, Whois-based features, DNS answer-based features; and Active Probing features. These five feature groups have a total of 14 features for detecting malicious DNS. The classification algorithm used in the paper is the J48 Decision Tree algorithm. For the APT attack detection technique based on network traffic, the authors presented 6 main features. After detecting an APT attack on both DNS log and Network traffic, the authors used a correlation analysis technique to detect which computer addresses in the system were infected with APT malware. However, in the paper, the authors didn't present details of this correlation calculation method.

In the paper [7], the authors combined the J48 Decision Tree algorithm with 4 main feature groups: DNS request and answer-based features; Domain-based features; Time-based features; and Whois-based features to detect APT malware command and control domains (C&C Domain). The Global Abnormal Forest and KNN machine learning algorithms are used in this study. The statistical correlation analysis technique is used by the authors to find out some of the new abnormal features of APT attacks. However, the authors didn't present data sources from which these abnormal features would be extracted.

In [8], the authors used 3 main groups of features to detect the domain APT, which are Domain name lexical features; Ranking features; DNS query features and Random Forest algorithm.

In the article [9], the author used the correlation analysis technique between DNS log and Network traffic, and some machine learning algorithms such as KNN, SVM to detect APT attacks.

Yan et al. [10] proposed the method of using the CNN deep learning algorithm to detect APT attacks based on DNS Activities. Accordingly, the authors extracted three main groups of features: Domain Name-based Features; Feature of the Relationship between DNS Request Behavior and Response Behavior; Feature of the Relationship between DNS Request Behavior and Response Behavior on a dataset of 4,907,147,146 pieces of initial data of 47 days DNS request records of Jilin University Education Network combined with CNN algorithm to detect APT attack behavior. There are also some other approaches for detecting malicious domains that support APT attack detection, including Vinayakumara et al. [11] used deep learning algorithms; and Nguyen [12] proposed using neutrosophic sets.

2.2 Detecting APT IP

Cho et al. [13] proposed an APT attack detection method based on the flow network using deep learning. Accordingly, in their research, the authors used some deep learning algorithms such as Multilayer Perceptron (MLP), Graph Convolutional Network (GCN), and BiLSTM-GCN model that combines Bidirectional Long Short-Term model Memory (BiLSTM) with GCN in order to analyze and re-represent information of APT attack IP based on network flow. The experimental results in the paper show that the BiLSTM-GCN combined deep learning model gives the best results on all metrics. We noticed that the approach of Cho et al. is very good and reasonable, but it requires a large and cumbersome computational system to implement.

In the study [14], Cho et al. proposed an APT attack detection method based on C&C servers using the Random Forest algorithms. In addition, the study [14] presented and listed several tools that assist in APT attack detection including Symantec, Forcepoint, McAfee, Kaspersky Lab, Fortinet, Cisco, Palo Alto Networks, and FireEye.

3 Proposing the APT Attack Detection Method Based on Network Traffic

3.1 Proposed Model of Detecting APT Attack Based on Network Traffic

The components of the APT attack detection model that we propose include:

- Detecting APT domain: The method of detection APT domains is responsible for classifying domains into clean or APT based on Network traffic using machine learning. To accomplish this task, we propose to use a combination of abnormal behavior analysis method and a classification method using machine learning. Details of abnormal behaviors will be presented in Section 3.2.1 of the paper. After identifying the suspicious domains as part of the APT attack campaign, we will use these domains as features for the APT IP detection process.
- Detecting APT IP: With the detection of malicious domains by machine learning techniques, it is possible to conclude that the system was hacked. Next, we will combine these detection results to assist in the APT IP detection process. To accomplish this task, Network traffic will first be evaluated and extracted abnormal behavior based on IP. We then use the Random Forest machine learning algorithms to classify these behaviors. Details on how to detect APT IP are presented in Section 3.3.1 of the paper.

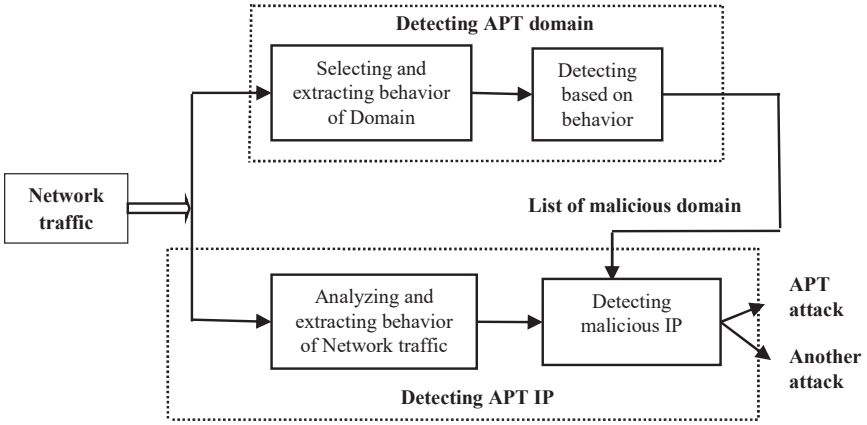

Figure 1 Proposed model of detecting APT attack based on Network Traffic using machine learning.

3.2 APT Domain Detection Method

3.2.1 Selecting and extracting features

To detect APT domains in Network Traffic, we extracted abnormal behaviors of these domains. In this study, in addition to the features described in the paper [8], we add 5 features extracted from the DNS log to detect abnormal behavior of DNS traffic. These are 5 new features that contain characteristics of APT attacks. The list of domain features that we choose and proposed is shown in Table 1 (new features are added symbols *).

The characteristics of the five new features that we propose in Table 1 are as follows:

- *The number of domains sharing the same IP:* in an attack APT, an individual attacker rarely owns more than 30 dynamic domains to point to the address of the C&C control server because it is not necessary and difficult to storage domains. Therefore, the number of domains that together point to the same IP address is usually less than 30.
- *The IP address in the same class B range of known C&C servers:* According to the results of statistical analysis of known C&C servers, there are many C&C servers having IPs in the same class B range together or even class A range. There can be 2 main reasons for this problem. First, the number of attackers who rent VPS virtual servers as C&C servers is increasing because the VPS server runs stably and is difficult to trace back and easy to manage. VPS servers rented from the same service provider will most likely have the same Class B IP addresses. Second, some advanced attackers have created a special system for C&C servers.
- *The daily similarity:* The daily recurring activities that we want to track for a domain address are the ones that change the IP address at certain times of the day. An APT attacker, for example, often points a domain to the C&C server's IP address at the start of business hours of the day and to meaningless IP addresses after the end of business hours. Some special malware connects to the C&C server at fixed times of the day. Tracking like that will be of great help to the detection process.
- *Same query number in the same time window:* This feature means that in the same period of time, the total number of DNS queries for that domain is equal. If the infected machine is going online but its connection to the C&C server is down for some reason, the infected machine sends a large number of DNS queries.

Table 1 List of domain features

No.	Group	Feature	Data Type	Description
1	Domain	Domain name length	Integer	The length of the domain
2	name lexical features (L)	Domain name token count	Integer	The number of tokens extracted from domain name separated by character '.'
3		Average domain token length	Real	Average length of the tokens.
4		Longest domain token length	Integer	Length of the longest token.
5		Number of IP address in domain name	Integer	Number of IP addresses in the domain name
6		Number of special characters	Integer	Number of special characters in the domain name
7		Number of digits	Integer	Number of digits in the domain name
8		Number of continuous digits	Integer	Number of continuous digits in the domain name
9		Longest continuous digits length	Integer	Length of the longest continuous digit series
10		Number of continuous letters	Integer	Number of the continuous letters in the domain
11		Longest continuous letters length	Integer	Length of the longest continuous letter series
12		Maximum Levenshtein ratio	Real	Maximum Levenshtein ratio compared to popular domain names
13		Brand name presence	Binary	Does the brand name exist in the domain name?
14	Ranking features (R)	Rank in Alexa	Integer	Rank of the domain name in the list of one million of common domain names from Alexa
15		Rank in Majestic	Integer	Rank of the domain name in the list of one million of common domain names from Majestic

(*Continued*)

Table 1 Continued

No.	Group	Feature	Data Type	Description
16		Rank in Domcop	Integer	Rank of the domain name in the list of ten million of common domain names from Domcop
17	DNS query features (D)	Resolved IP count	Integer	Number of IP addresses returned in the DNS query.
18		Distinct country count	Integer	Number of country from IP addresses
19		Silent IP ratio	Real	Ratio of inappropriate domains
20		HTTP response status	Integer	Status of the returned HTTP response
21		Name server count	Integer	Number of the name servers returned in the DNS query.
22		Name server IP count	Integer	Number of IP addresses of the name server in the DNS query
23		Name server Country count	Integer	Number of countries in which the DNS servers are located
24		**The number of domains sharing the same IP***	Integer	This feature is described in detail below
25		**The IP address in the same class B range of known C&C servers***	Binary	This feature is described in detail below
26		Mail exchange server count	Integer	Number of mail exchange servers returned in the DNS query
27		Time to live (TTL)	Integer	Time to live of the cached records for the domain name in the DNS server
28	Time value-based features (T)	**The daily similarity***	Binary	This feature is described in detail below
29		**Same query numbers in the same time window***	Binary	This feature is described in detail below
30		**Very low frequency query***	Binary	This feature is described in detail below

- *Very low frequency query:* Some advanced APT malware executes the query to the domain to determine the C&C control server with a very low frequency in order to avoid detection. The time between queries can up to several dates, even several weeks or months. This could be the behavior of sophisticated malware designed for the purpose of evading detection. According to the experiment, these domains usually have other common signs attached to them such as a web server that has complete content, stable IP and TTL addresses.

3.2.2 APT domain detection algorithm

As mentioned above, to detect APT domains, we will use the Random Forest algorithm. The Random Forest algorithm has been evaluated by many studies as a machine learning algorithm that has many outstanding advantages compared to other classification machine learning algorithms. The studies [15, 16] presented the theoretical details of the Random Forest algorithm. In this paper, we will apply this algorithm and change some of its parameters in order to evaluate its effectiveness for the APT domain classification process.

3.3 Development of Malicious IP Detection Module

3.3.1 Selecting and extracting abnormal behaviors of IP

In this paper, we propose 5 features that represent abnormal behavior of connections in Network Traffic. A list of these 5 features is shown in Table 2.

The characteristics of the five features representing abnormal connections of IP in Network Traffic in Table 2 are as follows:

- *Abnormal protocols and ports.* To pass the control of the firewall of the target network, malware often uses the common service port such as port 80, 8080, 443, 8000, etc. As you know, when connecting to the C&C server, the malware had to ask for operating system permission.

Table 2 List of features for detecting abnormal connection behavior of IP

No.	Feature	Data Type
1	Abnormal protocols and ports	Binary
2	A large discrepancy between the transferred and received data	Binary
3	Abnormal TCP connection	Binary
4	Heartbeat Traffic	Binary
5	Abnormal data fluctuation	Binary

Then, the server opens a random port (not a service port and usually a large port). Thus, with the problem of detecting APT attacks on the server, when the server opens a certain port that does not run properly the service, the server is likely infected with malware and malware is communicating with C&C Server. For example, detecting that on network traffic has HTTP traffic that does not go through port 80 or port 8080. To find abnormal ports that are not running the correct service according to specified internet standards, the first step is to find out the strange IPs that the server queried to in the DNS packet. From there, we will look for the queries that the server queries to that IP. From those records, extracting protocols and service ports where the server sends data out. Checking if they are suitable or not. Otherwise, they are abnormal abnormal protocols and ports. The reason for such the defining principle is that when a server provides an outgoing service with a specified port, it always listens to the request and returns the response through that port. For example, a web service with HTTP protocol has a default port 80. Thus, the webserver always listens to the request and returns the response over port 80.

- *A large discrepancy between the transferred and received data.* Usually, the data that the server transfer to the client is usually greater than or at least equal to the data sent by the client. But when the server is infected with malware, the data transferred to the client will be much higher than the data transferred from the client to the server. Because, when the server is infected with malware, the data transferred to the client contains files, and when the client sends data to the server, it is a search or download command. In order to find data transferred from the C&C server, we need to extract strange IPs from DNS records. From there, any connection that transfers data from that IP and sent packet size that is taken in the TCP_len field is larger than the size of the packet transferred from the server to that IP address, it violates this case.
- *Abnormal TCP connection.* When hackers send attack commands to the server that is infected with malware, the TCP connection time between the infected server and the C&C server is very long because the commands that a hacker sent are the commands to search or download the file. Therefore, the server will take a lot of time to find and return the response to the C&C server. Here, we consider connections that have the time greater than or equal to one minute as abnormal connections. To determine the connection time of host A to host B, we specify that

if the FIN flag of host A = 1 and the flag ACK = 1 and the SYN flag of host B = 0 and the flag ACK = 1, the connection time of a connection is equal to the subtraction result of time_epoch of 2 records. If the connection time is greater than or equal to 60 seconds, it is an abnormal TCP connection.

- *Heartbeat Traffic.* After the malware successfully infected, the malware sends packets to the C&C server to inform the C&C server that the malware has successfully infected and is still on the victim's computer. These packets are sent at a fixed time and the packet size is always the same. When a malware enters the victim's computer, the malware begins to connect to the C&C Server and sends packets of the same size in the same period of time to inform C&C Server that the malware is still online. This is called Heartbeat Traffic. To determine Heartbeat Traffic, we will take the values of the TCP_len and time_epoch fields to determine the size and time of the packet. From there, check the size of the packets in periods of time such as dates, months, and years (depending on the request) to determine heartbeat.
- *Abnormal data fluctuation.* The data of C&C servers are often small and steadily but when hackers start sending data from the infected server to C & C server, the data will increase dramatically. As with the previous signs, the first step identifying strange IPs in DNS records. Then, get the value of the TCP_len field of all records whose destination IP is a strange IP. From there, find the maximum packet size and calculate the average size of the packets. If the size of the largest packet is greater than 60% of the average size of packets, that data is abnormal data.

3.3.2 APT IP detection algorithm

To detect APT IP, we will also use RF algorithm as APT domain detection process.

4 Experiments and Evaluation

4.1 Installation Requirements

The system has a specific configuration and environment as follows:

- Software requirements: Python version 3.6; Spark version 2.3.0; Hadoop version 2.7; Java (JDK) 8; ID Bro version 2.5.3; Ubuntu 16.04.4
- Hardware requirements: RAM 8GB; CPU Intel Core i5 3.50GHz, 4 Cores, 6th Generation.

4.2 Training and Testing Dataset

4.2.1 Training dataset

(a) The training dataset for detecting malicious domain

Experimental dataset in this paper consists of 164077 domains that are collected at [18–22]:

- Dataset of Benign domains: This dataset consists of 79910 benign domains that have been collected from the most well-known domain names on the Internet.
- Dataset of unknown and malicious domains: This dataset covers 84167 phishing domains derived from Phish Tank, C&C domains, Malicious domains list.

(b) The training dataset for detecting malicious IP

- 28 pcap files of APT attack malware were filtered from the CTU Malware Botnet dataset [23–25]. The dataset includes 985.595 DNS queries; 50 domain names consisting of 26 domain names related to APT attacks, and 24 clean domain names; 921 IP addresses, 581 public IP addresses, 71 IP addresses related to APT attacks.
- 33 pcap files of APT attack malware from APT – Mila dataset [26]. The dataset includes 272 DNS queries; 45 domain names consisting of 18 domain names related to APT attacks, and 27 clean domain names; 105 IP addresses, 75 public IP addresses, 25 IP addresses related to APT attacks.

4.2.2 Testing dataset

(a) The testing dataset for detecting APT domain

To test the APT domain detection model, we build a malicious domain detection model based on the dataset presented in part (a) of Section 4.2.1. To evaluate the effectiveness of the proposed method in the paper, we will conduct experiments and compare the results of our research with some other studies [8, 17]. Then, based on the results built from the training process, we will use this training model to test its ability to detect with the DNS log dataset containing the APT domains presented in part (b) of Section 4.2.1.

(b) The testing dataset for detecting APT IP

Based on the dataset presented in part (b) of Section 4.2.1, we randomly divide this dataset into 2 parts including 80% for training and the remaining 20% for testing.

4.2.3 Experimental scenario and evaluation criteria

(a) Experimental scenario

In this paper, to evaluate the effectiveness of each model, we divide the experimental process into 3 different scenarios as follows:

- Scenario 1: Detecting APT attack domains.
- Scenario 2: Detecting APT IP without using domain features.
- Scenario 3: Detecting APT IP using domain features.

(b) Evaluation criteria

To evaluate the performance of the APT attack detection system, 4 different measures are used including accuracy, precision, recall, and f1-score. These metrics are calculated based on the following components:

- True positive (TP) is the number of malicious IPs or domains correctly classified.
- True negative (TN) is the number of normal IPs or domains correctly classified.
- False positive (FP) is the number of normal IPs or domains missed classified into malicious or APT.
- False negative (FN) is the number of malicious IPs or domains missed classified into normal.

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \tag{1}$$

$$\text{precision} = \frac{TP}{TP + FP} \times 100\% \tag{2}$$

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\% \tag{3}$$

$$\text{F1-score} = \frac{2 \times \text{precision} \times \text{Recall}}{\text{precision} + \text{Recall}} \tag{4}$$

4.3 Some Experimental Results

4.3.1 Malicious domain detection results

(a) APT domain classification results

The experimental results in Table 3 show that when using the Random Forest algorithm to detect the APT domain that we propose, all measures are very high, namely acc: 97.56%, precision: 98.74%, recall: 96.48%, F1-score: 97.60%. Especially, the FNR rate is significantly lower. This proves

Table 3 Results of training the model of malicious domain detection

Algorithm	Accuracy %	Precision %	Recall %	F1 %	FPR %	TNR %	FNR %
RF [8]	96.16	96.70	95.79	96.24	3.44	96.56	4.21
RF [17]	96.92	97.18	96.80	96.99	2.95	97.05	3.20
RF [our proposal]	97.56	98.74	96.48	97.60	1.30	98.70	3.52

that this algorithm has a low rate of false prediction from benign to malicious domains. Besides, when comparing the results of detecting malicious domain, we notice that our approach is better than some other approaches. This shows that the selected and extracted features of malicious domains presented the clear difference between the malicious domains and the clean domains. Next, we will use this malicious domain detection model to test the dataset of actual APT attack domains.

(b) Testing the model of APT domain detection

With the test dataset of DNS logs of APT domains as shown in Section 4.2.1, we have the test results of the APT domain detection model shown in Table 4 below.

Table 4 Test results of the malicious domain detection model

Accuracy %	Precision %	Recall %	F1 %	FPR %	TNR %	FNR %	Testing Time (s)
92.54	97.89	88.31	92.85	0.61	98.39	11.69	0.091

The results presented in Table 4 shows that the accuracy of the method of detecting APT attacks by DNS behavior is relatively good (92.54%). The false detection rate FPR is low (0.61%). Can see that the process of testing the malicious domain detection model gave really good results even though the experimental dataset has a difference in the proportion of normal domains and APT domains. This result shows that malicious domain features that were selected during the training process have relatively accurately and fully defined the behavior of the APT domain. With this result, the paper has provided detection systems for malicious domains in general and APT domains in particular a novel method to detect and classify domains.

4.3.2 Experimental results of the APT IP detection

(a) Experimental results of the APT IP detection without domain feature

To evaluate the effectiveness of the APT IP detection process, in this paper, we will use the Random Forest algorithm with a change in the number of

Table 5 Results comparing methods of detecting APT IP without domain feature

Algorithm	Accuracy (%)	Precision (%)	Recall %	F1 (%)	FPR %	TNR %	FNR %
RF [50 trees]	**93.12**	**73.44**	**86.55**	**79.46**	**5.69**	**94.31**	**13.45**
RF [30 trees]	91.97	69.87	84.05	76.31	6.59	93.41	15.95
RF [100 trees]	92.35	71.0	84.95	77.35	6.31	93.69	15.05

Table 6 Results comparing methods of detecting APT IP using domain feature

Algorithm	Accuracy (%)	Precision (%)	Recall %	F1 (%)	FPR (%)	TNR %	FNR %
RF [50 trees]	**94.62**	**80.69**	**85.45**	**83.0**	**3.72**	**96.28**	**14.55**
RF [30 trees]	93.01	76.60	78.55	77.56	4.36	95.64	21.45
RF [100 trees]	93.98	78.55	83.7	81.04	4.15	95.85	16.3

trees. Table 5 below shows the results of APT IP detection using the Random Forest algorithm with the features defined in Table 2.

Through the experimental results in Table 5, we notice that when the domain feature is not used, the Random Forest algorithm with 50 trees gives the highest accuracy with the accuracy of 93.12%. If only evaluating based on this accuracy, the system classified very well. However, if looking at the precision (the ratio of the number of correctly predicted APT IP among those classified as APT IP) of all is low and the FPR rate is quite high. The cause of this problem is that the test dataset has a big discrepancy between the proportion of clean data and malicious data. Besides, the selected and extracted features contain many characteristics of APT attacks, but these characteristics do not appear much in the test dataset.

(b) Experimental results of the APT IP detection using domain feature

Table 6 shows the results of APT IP detection when using domain features. Obviously, when using domain features, the Accuracy, Precision, and FPR measures on all 3 experiments significantly improved compared to without using domain features. The Random Forest algorithm with 50 trees still gave the best results with Accuracy increased from 93.12% to 94.62%, Precision increased from 73.44% to 80.69% and false detection rate FPR decreased from 5.69% to 3.72%. The cause of this problem is that in reality, APT attacks often use many different methods, tools, and means to attack the same target, so if monitoring systems have the ability to analyze and calculate the relationship between domain and IP, this will open many bases to conclude about signs of APT attacks in the system. This is a very good result for APT

IP detection when the experimental dataset has a big discrepancy. However, in practice, such datasets correctly represent the rules of network monitoring systems. The experimental results in Table 6 shows that the APT IP detection approach using domain correlation results gives completely good and outstanding results compared without using domain features. In addition, we have demonstrated that for the process of detecting APT domains, instead of having to find and define specific characteristics and features of APT attacks, monitoring systems can use the combination of features and the correlation between the components in the dataset.

5 Conclusion and Future Development Direction

APT attacks have been and will always be major challenges for information security systems. In this paper, in order to optimize the APT attack detection process, we propose a new approach based on the analysis and evaluation of Network Traffic components using machine learning. Accordingly, we have proposed a method to detect APT domain based on the characteristics and behavior of Network traffic using machine learning. Experimental results have shown that our proposed method has yielded better results than other studies. This shows that the method of selecting and extracting features and the behavior of the domain presented the clear difference between APT domains and clean domains. Besides, based on the abnormal behavior of Network traffic, we propose a method to detect APT IPs using machine learning. The experimental results on the application of the Random Forest machine learning algorithms for the classification process of APT IP and clean IP brought relatively good results though the experimental dataset has a huge difference between the proportion of clean IPs and APT IPs. In addition, the results of a number of scenarios for APT IP detection prove that our proposed method that combines the behavior and features of the IP and the domain has better results than individual detection methods. This result proves that our proposal is not only correct and scientific but also provides various mechanisms for systems that monitor and detect cyber-attacks in general and APT attacks in particular. In the future, we will expand the scope and research direction of detecting APT attacks not only on two components, namely domain and IP, but also on other components such as HTTTP, TLS, Flow, etc. These are very important components. Building and synthesizing behavior profiles of these components will help the detection system recognize and evaluate the correlation between components in order to supplement the mechanisms for APT detection.

References

[1] Quintero Bonilla, Santiago & Rey, Ángel. A New Proposal on the Advanced Persistent Threat: A Survey. Applied Sciences. 2020, 10(11), pp. 38–74.

[2] Adel, A., Ankur, C., Sowmya, M., Dijiang Huang, H.: A Survey on Advanced Persistent Threats: Techniques, Solutions, Challenges, and Research Opportunities. IEEE Commu. Sur. & Tu. PP99(1–1), 1–29 (2019).

[3] Zimba, Aaron, Chen, Hongsong, Wang, Zhaoshun, Chishimba, Mumbi. Modeling and detection of the multi-stages of Advanced Persistent Threats attacks based on semi-supervised learning and complex networks characteristics. Future Generation Computer Systems. Volume 106, 2020, pp. 501–517.

[4] Sadegh, M.M., Rigel, G.J., Birhanu, E., Ramachandran, S., HOLMES: Real-time APT Detection through Correlation of Suspicious Information Flows. In: 2019 IEEE Symposium on Security and Privacy, pp. 1137–1152, San Francisco, CA, USA, 19–23 May 2019.

[5] Lajevardi, Amir, Amini, Morteza. A semantic-based correlation approach for detecting hybrid and low-level APTs. Future Generation Computer Systems. Vol. 96, 2019, pp. 64–88.

[6] Weina, N., Xiaosong, Z., GuoWu, Y., Jianan, Z., Zhongwei, R., Identifying APT Malware Domain Based on Mobile DNS Logging. Mat. Pro. in. Eng. 2, 1–9 (2017).

[7] Zhao, G., Xu, K., Xu, L., Wu, B., Detecting APT malware infections based on malicious DNS and traffic analysis. IEEE Access. 3, 1132–1142 (2015).

[8] Do Xuan Cho, Ha Hai Nam. A Method of Monitoring and Detecting APT Attacks Based on Unknown Domains. Pro. Com. Sci. 150, 316–323 (2019).

[9] Jiazhong Lu, Kai Chen, Zhongliu Zhuo, XiaoSong Zhang. A temporal correlation and traffic analysis approach for APT attacks detection. Cluster Computing (2017). pp. 1–12.

[10] Guanghua Yan, Qiang Li, Dong Guo, Xiangyu Meng. Discovering Suspicious APT Behaviors by Analyzing DNS Activities. Sensors 2020, 20, 731; doi:10.3390/s20030731.

[11] R. Vinayakumara, K.P. Somana, P. Poornachandranb. Detecting malicious domain names using deep learning approaches at scale. Journal of Intelligent and Fuzzy Systems. 2018, 34, 1355–1367.

[12] Van Can, Nguyen et al. A New Method to Classify Malicious Domain Name Using Neutrosophic Sets in DGA Botnet Detection. Journal of Intelligent and Fuzzy Systems. 2020, 36, 4223–4236.
[13] Cho Do Xuan, Hoang Mai Dao, Hoa Dinh Nguyen. APT attack detection based on flow network analysis techniques using deep learning. Journal of Intelligent & Fuzzy Systems, vol. 39, no. 3, pp. 4785–4801, 2020.
[14] Cho Do Xuan, Lai Van Duong and Tisenko Victor Nikolaevich, "Detecting C&C Server in the APT Attack based on Network Traffic using Machine Learning", International Journal of Advanced Computer Science and Applications(IJACSA), 11(5), 2020. http://dx.doi.org/10.14569/IJACSA.2020.0110504.
[15] Shai, S.S., Shai B.D., Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press (2014).
[16] Leo, B., Random Forests. Ma. Lear. 45(1), 5–32 (2001).
[17] Xuan, Cho. Malicious domain detection based on DNS query using Machine Learning. International Journal of Emerging Trends in Engineering Research. No. 8, 2020, pp. 1809–1814.7.
[18] OpenDNS public domain lists of domain names for training/testing classifier. https://github.com/opendns/public-domain-lists [access date 1/4/3018].
[19] Malware Domain List. http://www.malwaredomainlist.com/ [access date 1/4/2020].
[20] Join the fight against phishing. https://www.phishtank.com/ [access date 1/4/2020].
[21] Alexa – Top Sites for Countries. https://www.alexa.com/topsites/countries [access date 1/4/2020].
[22] Public-domain-lists. https://github.com/opendns/public-domain-lists [access date 3/4/2020].
[23] APTNotes – Github Repo. https://github.com/kbandla/APTnotes [access date 3/4/2020].
[24] APTNotes – Website. https://aptnotes.malwareconfig.com/Targeted [access date 3/4/2020].
[25] Cyber Attacks Logbook (Kaspersky). https://apt.securelist.com/ [access date 3/4/2020].
[26] DeepEnd Research: List of malware pcaps, samples, and indicators for the Library of Malware Traffic Patterns. https://contagiodump.blogspot.com/2013/08/deepend-research-list-of-malware-pcaps.html [access date 3/4/2020].

Biography

Cho Do Xuan is currently a lecturer at the Faculty of Information Technology at Posts and Telecommunications Institute of Technology and FPTU in Vietnam In 2008, received a bachelor's degree in the Saint Petersburg Electrotechnical University "LETI" on a specialty "Computer science and computer facilities", Russia. In 2010, graduated a masters from the Saint Petersburg Electrotechnical University "LETI" on a specialty "Computer science and computer facilities", Russia. In 2013, received a PhD in the Saint Petersburg Electrotechnical University "LETI", on a specialty CAD. Russia. Area of scientific interests – modeling, control systems, algorithmization, information security.

E-mail: chodx@ptit.edu.vn and chodx@fe.edu.vn