Mitigating Advanced Persistent Threats Using A Combined Static-Rule And Machine Learning-Based Technique

Oluwasegun Adelaiye
Department of Computer Science
Bingham University
Karu, Nigeria
oluwasegun.adelaiye@binghamuni.edu.ng

Aminat Ajibola
Department of Computer Science
University of Abuja
Abuja, Nigeria
aminat.ajibola@uniabuja.edu.ng

Abstract—**Advanced Persistent Threat is a targeted attack method used to maintain undetected unauthorized access over an extended period to exfiltrate valuable data. The inability of traditional methods in mitigating this attack is a major problem, which poses huge threats to organizations. This paper proposes the combined use of pattern recognition and machine learning based techniques in militating the attack. Using basic statistical test approach, a dataset containing 1,047,908 PCAP instances is analyzed and results show patterns exist in identifying between malicious data traffic and normal data traffic. The machine learning on the other hand, is evaluated using three algorithms successfully: KNN, Decision Tree and Random Forest. All algorithms showed very high accuracies in correctly classifying the data traffic. Using the algorithm with the highest accuracy, Random Forest is optimized for better effectiveness.**

Index Terms—**Information Security, Traffic analysis, Intrusion Detection, zero-day, Packet capture**

I. Introduction

The risk posed by a successful information security breach has increased rapidly in recent times and does not just affect machines but also risks human wellbeing and existence [1]. The cost of a successful attack is estimated at 7.2million dollars per organization [2]. A recently discovered attack technique termed Advanced Persistent Threat (APT) named based on its attack technique, is a type of attack defined by the National Institute for Science and Technology (NIST) as a highly-skilled expert with notable resources who aims at creating and expanding control within an organization's Information Technology (IT) substructure to obtain confidential information, deny or negatively affect censorious programs and missions, or create a platform to aid future attacks. [3], [4]

APT being a multi-plane attack but majorly a compound network attack rapidly evolves and spreads while continuously changing its infiltration techniques, posing a great threat to organizations. Fueled with the increased growth in computer-based solutions and networked communities, this targeted threat has been drawing increasing attention among security experts.

Advanced Persistent Threats (APT) in the first half of 2011, gained prominence [6] through the occurrence of several high profile and persistent information security breaches reported by large global organizations including the military, financial, energy, nuclear, education, aerospace, telecom, chemical, and government sectors. Red October, Operation Aurora, RAS breach operations, Duqu, Ke3chang operation, Flame, Stuxnet, Snow Man, and Mini Duke are a few popular Advanced Persistent Threat (APT) attacks that occurred in 2011 [7], [8]. APTs are often associated with cyber-espionage activities, aiming to steal highly confidential information which includes trade secrets, Intellectual Property, national security data etc., for monetary gain or geared towards the sabotage of strategic infrastructures [8].

Most of the earlier contributions on APTs were based on comprehensive studies of earlier APT attacks that attempted to identify the inherent characteristics of these attacks, propose the APT attack stage model and highlight some generic countermeasures to mitigate APTs [9] [10] [11].

In this respect, this work proposes the combination of static rule and machine learning anomaly detection based techniques in thwarting APT attacks thereby providing new mechanisms for detecting and preventing such forms of information security breaches. Having had an introduction to APT the next section presents problems responsible for the build of this research.

II. Statement of Problem

From the successes of Advanced Persistent Threat (APT) attacks that have occurred and the results of recent studies, it is evident that there exists a challenge in detecting Advanced Persistent Threats. The seriousness of APT is visible from the high profile attacks and exfiltration of data in the attacks on sensitive organizations like Sony, Citigroup, RSA security, NASA, FBI, Fox broadcasting etc.,[6] These organizations had traditional security methods implemented but yet could not mitigate the attack. Researchers have identified and looked into this problem, these largely relate to the inefficiency of traditional prevention and detection techniques in mitigating targeted attacks. [12] [13]

978-1-7281-5160-1/19/$31.00 ©2019 IEEE

The ineffectiveness of traditional mitigation techniques in preventing against Advanced Persistent Threats (APT) has led to the loss of valuable data by large organizations and government agencies. Most of the methods that have been created have not been effective in detecting and prevent APT activities in the user, application, network or physical plane.

Despite the improvements in defending and protecting against security breaches, the ability of APT attacks to bypass security mechanisms shows that the threat still exists.

In this respect, this paper through the use of multiple anomaly detection techniques using static rule-based detection and adopting an optimized Ensemble learning algorithm aims to implement and assess the effectiveness of the proposed security mechanisms in thwarting APT attacks. This approach uses PCAP files, which are easily extractible during packet transmission.

III. Mitigating Advanced Persistent Threats Effects

Researchers have proposed different approaches to mitigating APT. Adelaiye et al. [1] reviewed the approach of 25 researchers. The results of their work showed disparate degrees of effectiveness. The frequencies of the methods employed by these researchers are shown in Fig. 1. The utilization of 12 mitigation techniques by 25 researchers in mitigating Advanced Persistent Threats is presented in Fig. 1, out of which 6 of the 25 researchers proposed the use of anomaly detection in mitigating the threat and 5 researchers proposed a combined implementation with traffic/data analysis. The data and the results of the data and traffic analysis are used in identifying and detecting normal behavior versus abnormal behavior. Lamprakis et al. [14] and Friedberg et al. [15] also proposed the use of whitelist, which combines three methods to militate the threat. The use of blacklist has been said to be ineffective in mitigating targeted and sophisticated attacks as it works based on pre-identified malware which is the reason for adopting the use of whitelist in detecting the presence of malicious activity as in [14] and [15].

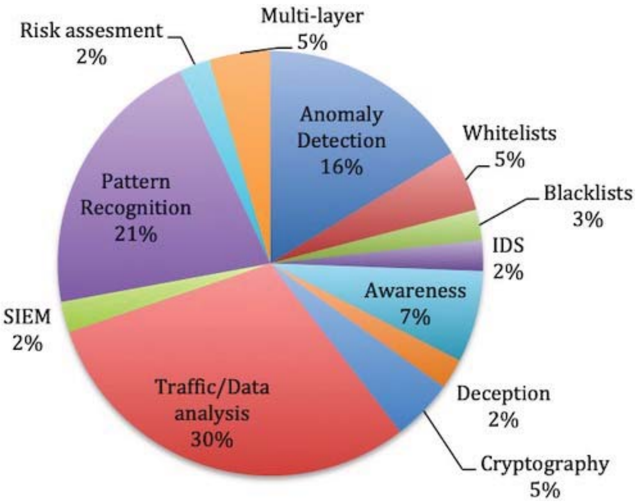

Fig. 1. Pie chart showing Mitigation Techniques Used against APT by 25 researchers [1]

A gene-based approach similar to traffic/data analysis in detecting Advanced Persistent Threats was employed as in [18] who identified some similarities with APT attacks using the pattern of pre-existing attacks that have occurred and combined this approach with anomaly detection as seen in Table 1. Other similar approaches to mitigating APT are shown in Table 2.

TABLE I
Implementation of Anomaly Detection

Author	*Anomaly Detection*	*Whitelist*	*Traffic/ data Analysis*	*Pattern Recognition*	*No of Methods*
Lamprakis, et al. [14]	√	√	√		3
Friedberg et al. [15]	√	√	√		3
Skopik et al. [16]	√		√		2
Vance [17]	√		√		2
De Vries et al. [12]	√		√		2
Wang et al. [18]	√			√	2

As seen in Table 2, Ghafir et al. [19] used machine learning correlation analysis. The machine learning collects the output of detection methods to efficiently classify APT alerts. This study showed 84.8% accuracy. Chandran et al. [20] achieved an accuracy level of 99.8% using random forest algorithm in predicting the occurrence of APT. Support Vector Machine (SVM) algorithm was also used on 1228 extracted log events and showed a 98.67% accuracy level [21]. Different algorithms have been applied to mitigating advanced persistent threats, which are majorly: SVM, K-Nearest Neighbor (KNN), Decision Tree and Random forest [12] [19] [22] [23].

TABLE II
Related Works

Author	*Method*	*Accuracy*
Ghafir et al. [19]	Machine Learning correlation analysis	84.8%
Chandran et al. [20]	Random Forest	99.8%
Schindler [21]	Simple Vector Machine	98.6%

The choice of combining two methods in mitigating APT is based on the recent approaches by researchers citing improved accuracy and consistency, and also in an attempt to reduce the chances of false positives.

Having presented mitigation approaches for Advanced Persistent Threats and related works in mitigating APTs, the next section presents the methodology employed in providing an improved mitigation approach.

IV. Material and Methods

In meeting up with the objectives of the research and proposed techniques for mitigating APT, the method to be used is broken down into three parts: 1. Static Rule-Based Anomaly Detection (Statistical Analysis). 2. Machine Learning Based Anomaly Detection 3. Model Development and integration.

These methods fit into the major chain of activities in meeting up with the objectives of this research as shown in Fig. 2.

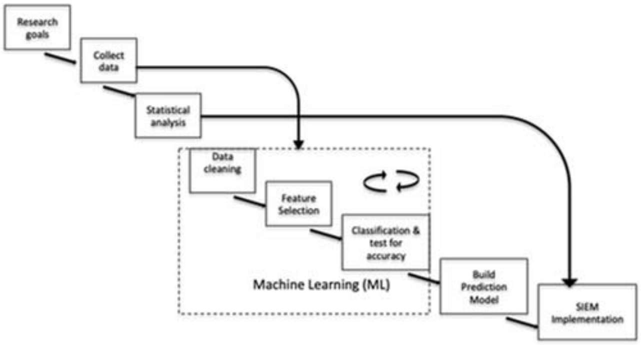

Fig. 2. Mitigation Plan from research goal to implementation

Fig. 2 illustrates the research plan for this study. The first part identifying research goals, data collection stage, which provides data, samples for both the statistical analysis stage a method used in static rule-based anomaly detection and the machine learning group stage. The machine learning group stage consists of three recursive stages that perform the learning operation using four algorithms in an attempt to select the best based on accuracy and speed. The output of this is fed to the development of an enhanced prediction model. The final stage combines both results static rule and the enhanced production model and implements it as an SIEM module.

A. Data Collection

Nigeria is yet to experience APT like attacks and sourcing a local dataset unique to this part of the world was not possible. A data set is gotten from Coburg University Germany, which provides fields of data carrying basic network traffic information as seen in Table 3. This dataset was built through monitoring network traffic of a business organization for 1 week using open stack. This dataset called Coburg Intrusion Detection Dataset (CIDD) consists of over 1 million records and consists of 11 fields (see Table 3) [24].

The CIDDS-001 data set is based on unidirectional data flow and for anomaly detection based research towards mitigating breaches. The data was collected using OpenStack in a business environment, which consists of multiple clients and servers. OpenStack is a cloud-based service it is an efficient and flexible resource management with features for network monitoring and packet capture [25]. The class column contains three distinct values: normal, victim and attacker. These values are allocated based on the direction of the traffic at that instance.

TABLE III
Attributes of CIDD dataset

Fields	*Description*
Duration	Duration of traffic flow
Protocol	Transport Protocol Used
Source IP address	Source IP address
Destination port No.	Destination Port
Destination IP address	Destination IP address
Packets	No. of Packets transmitted
Bytes	No. of bytes transmitted
TOS	Type of Sevice
Class	Classification (Normal, Attacker and Victim)
Attack Type	Attack vector used
Attack ID	Unique identification for each attack vector type
Attack Description	Details about the attack parameters

The next section describes the statistical analysis in an attempt to build a static rule-based anomaly detection.

B. Static-Rule Anomaly Detection

Static rule-based anomaly detection uses finite sets of rules to detect anomalies. The algorithm presents the step-by-step procedure in the detection process using static rule approach:

```
Input $V = (a_i, b_i)$     [Indices affected by the rule applied]
Output: anomalous traffic,
$P = \{p_{1,} p_2, \ldots, p_n\}$   [captured data traffic, T all data traffic within network]
Begin
Initialize $P = \{\ \} \in T$ ,  $S = \{S_1, S_2\}$ where $S_1$ and $S_2$ are the rules threshold
For each $p_i$ where $\{a, b\} \subseteq P$
    $V \leftarrow \{a_i, b_i\}$
For each V (<, > or =) $S_1$ and V (<, > or =) $S_2$      [$S_1$ and $S_2$ are predefined rules]
  If $D \leftarrow a_i = b_i$
      Return D
End
```

The process of detecting intrusion using static rule-based anomaly detection can be logically explained using an algorithm to show a step-by-step procedure, this is shown above. The input data V, which is sourced from the traffic flow when the sensors capturing the traffic is active is the input data (a_i,b_i) required in detecting anomalies in that instance. The output returned is the identity of the anomalous traffic. P={ p_1, p_2,. . . ,p_n} refers to the traffic being captured a subset of all traffic within the network. The process begins by initializing the captured traffic P to empty and the rules S_1 and S_2 are set to their respective suspicious limit values. The two values a,b of each instance contained in the metadata of the traffic. A condition to selecting the abnormal traffic $V > S_1$ and $V < S_2$ is checked and stores the suspected anomalies in D← a_i=b_i. D is returned as abnormal traffic, which triggers flags to alert the administrator. The condition for the implementation of the static rule-based anomaly detection algorithm, as shown above, is through statistical analysis test for association. In finding patterns in traffic data a research hypothesis is adopted.

Research Hypothesis

H_0: There is no difference between the behavioral pattern of normal and malicious traffic with respect to source port, destination port, packets and bytes.

H_1: There is a difference between the behavioral pattern of normal and malicious traffic with respect to source port, destination port, packets and bytes.

C. Machine Learning-Based Anomaly Detection

Having utilized the Static rule-based anomaly detection technique a more advanced method is to be used to try and improve accuracy and to overcome unforeseen challenges. This method uses methods that can be likened to the behavior of a human being who is learning about something new to be able to respond to an event that occurs. There are four algorithms to be tested and evaluated to make sure the result has the highest level of achievable accuracy in classifying data and events. These algorithms are Ensemble, Nearest Neighbor, Decision tree and Simple Vector Machine. The algorithms use different methods of marching events based on similarities. The selection of the algorithms was based on their utilization in Advanced Persistent Threat like researches. [12] [19] [23] [26]

V. Results

In getting a finite set of patterns to implement static rule-based anomaly detection, Kruskal Wallis a nonparametric test done by comparing k independent samples is used after a test for normality is done. The results are represented in Table 4.

TABLE IV
Test for association

Field	Median (Average Rank Z)			*Test Statistic-H*	*P-Value*
	Normal	*Malicious (Attacker)*	*Malicious (Victim)*		
Source Port	8082 (-27.27)	51357 (50.56)	2701 (-18.87)	2941.21	0.000
Packets	1.0 (98.14)	1.0 (-79.17)	1.0 (-63.76)	11357.29	0.000
Bytes	120 (173.47)	58 (-124.25)	54 (-121.05)	31357.13	0.000

The results from Table 4, shows that for source port we can certainly infer that the median port number for normal and victim traffic is significantly lower than the median port number for attack traffic. This is based on the average rank (Z) for normal and victim is significantly lower than the overall mean rank (Z=-27.27 <− 1.96 & Z=-18.87<− 1.96), and the average rank for attacker is significantly higher (Z=50.56>1.96) than the overall mean rank. The p-value also shows enough evidence of a difference similar with packets and bytes. With the average rank (Z) for normal significantly higher than the overall mean rank (Z=-98.14>1.96), and the average rank for malicious traffic both for victim and attacker significantly lower (Z=-63.76< −1.96 & Z=-79.17< −1.96) than the overall mean rank, We can certainly infer that the median number of packets for normal traffic is significantly higher than the median number of packets for malicious traffic. We can also infer that the median size in bytes used in normal traffic is significantly higher than the median bytes used for malicious traffic based on the average rank (Z) for normal significantly higher than the overall mean rank (Z=173.47>1.96), and the average rank for malicious traffic significantly lower (Z=-124.25< −1.96 & Z=-121.05< −1.96) than the overall mean rank.

Having highlighted the results from the tests and obtained acceptable results the next section presents the machine leaning approach in creating an Ensemble detection technique.

A. Machine Learning-Based Appraoch

The machine learning approach to mitigating APT and other security threats is based on the fact that acquired knowledge gotten from experiences is vital in detecting malicious activity without human interference or control.

1) Data Cleaning: A dirty dataset or dataset with impurities needs cleaning for better results. On examining the dataset, the Bytes column shows that the dataset contains unexpected values. On inspecting the data set through sorting the data, it showed that the unexpected values were 0.06% of the whole dataset and so it would be safer to use the delete affected column method rather than replacing values which might introduce some amount of bias to the prediction model.

The data set does not just contain unexpected values but also unexpected data types. This affects the column, which is meant to be integer presented as a string. Other things looked at were unexpected values, consistency, type conversion, uniformity and verifying correctness. Things not considered are precision of data, duplicates, syntax errors, standardization (to be done later), scaling and transformation, Normalization (mostly with statistical methods), cross dataset errors and Outliers.

2) Feature Selection: For better results the noise data present needs to be identified to reduce the chances of misleading the model from achieving high accuracies. This method selects a subset of features present in the dataset based on some criteria. Using three methods:

1)Univariate Feature Selection Method
2)Feature Importance
3)Correlation Matrix

1) Univariate Feature Selection Method

This method uses statistical calculations in selecting non-negative features. Using Chi^2 test, the selected features are ranked from the best feature to the least feature. From the results, Bytes is the best feature and flows the least best feature. This is evident in the Chi^2 scores of 2.092691e+08 for Bytes and 0.000000e+00 for flows, others in-between are source port number (Src Pt), destination port number (Dst Pt), Type of Service (Tos), Packets and Duration.

2) Feature Importance

Using Extra Tree Classifier on the dataset, the feature importance results are presented in Fig. 3.

The graph results defer slightly from the chi test resul by ranking packets higher than bytes but are similar in

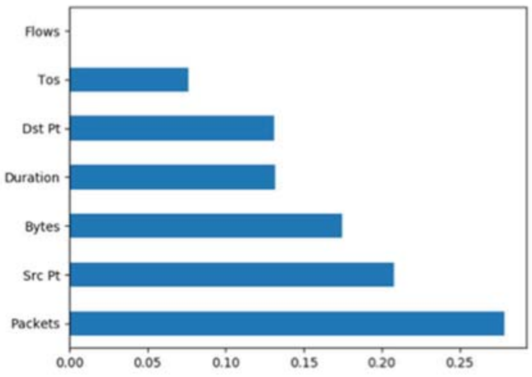

Fig. 3. Graph showing Feature Importance

identifying flows as the least important feature of the dataset. In between packets and flows are the other features where source port ranks second highest preceded by bytes.

3) Correlation Matrix

This approach correlates features using an x by x matrix where x is the number as well as represents each feature. The coefficient matrix shows the relationship between the selected features of importance using matched cells. Each feature correlates highly with itself as expected. Showing greater signs are bytes, packets and duration. The correlation between destination port number and Type of Service is the highest with a coefficient of 0.5 while the lowest is between destination port and source port with a coefficient of 0.92. "Flows" shows no correlation with any other feature including itself. Others with positive correlation include bytes and duration, packets and duration, and bytes and packets.

From all the tests done it is evident that "flows" is of no positive significance in classification and should be discarded.

3) Classification: Having cleaned the data and narrowed down the features, the classification of the data and testing for accuracy is next. The learning style to be used is the supervised methods.

TABLE V
MACHINE LEARNING APROACH

Algorithm	***Accuracy***	***Time Taken(Sec)***
KNN	99.74%	693.34
Support Vector Machine	87.11%	395.95(Datasize 5,240)
Decision Tree (CART)	99.84%	30.56
Random Forest	99.90%	78.95
Optimized Random Forest	99.95%	(Training Dataset size 70,000)

From Table 5, K Nearest Neighbor algorithm when applied to the dataset, the classification accuracy is 99.74%. Support Vector Machine (SVM) was unable to classify the whole dataset due to the size and but showed an accuracy of 87.11 using 5,420 instances. On using Decision tree algorithm, the result shows an accuracy level of 99.84%. The result with the highest accuracy is Random Forest, an ensemble approach to Decision Tree giving 99.90% accuracy in correctly classifying traffic data.

The Random Forest Algorithm having the highest accuracy is optimized. The modifications include load balancing and modification of the formula for the calculation of entropy. The Load balance approach randomly selects equal sets of data based on labels and feeds to the Random Forest algorithm. The calculation for entropy, which is modified as, indicated in (2) .

$$entropy -= p * log_2(p) \tag{1}$$

A constant 10 is introduced modifying the formula to

$$entropy -= p * 10 * log_2(p) \tag{2}$$

These modifications increased the accuracy from 99.90% to 99.95% using a much smaller dataset.

VI. DISCUSSION

Mitigating Advanced Persistent Threats have been an issue over time. From the results recorded in this study, we see great levels of accuracy and effectiveness in mitigating APT attacks. The Dataset used provides PCAP files for normal, attacker and victim traffic using multiple attack vectors. The dataset following a hybrid approach is tested for patterns using statistical methods for the test of association based on the labels. Findings show that most malicious traffic made up of both the attacker and victim instances, are mostly small in size using bytes as identified in Table 4. Table 4 also shows that the port numbers for malicious traffic are mostly from the private/dynamic port numbers. These patterns are finite and meet the conditions for static rule-based anomaly detection. From the algorithm for static based anomaly detection these conditions can easily be implemented to replace S_1 and S_2 which is used to filter the traffic and detect malicious activities. The idea is to utilize two methods in improving accuracy and reducing false positives. This method combined with Machine learning-based anomaly detection techniques shows great potentials in efficiently mitigating APT. The machine learning approach uses multiple algorithms to check for the most accurate. These algorithms selected based on similar studies done showed good signs in detecting APT exploits with the accuracy of 99.74%, 87.11%, 99.84% and 99.90% for K-Nearest Neighbor, Simple Vector Machine, Decision Tree and Radom Forest algorithm respectively. Selecting the algorithm with the highest (Random Forest) and optimizing it provides higher accuracies than earlier recorded. Using load balancing and modifying the formula used in calculating for best split called entropy improves the accuracy from 99.90% to 99.95% hence increasing the chances of correctly detecting the presence of malicious traffic. With 99.95% accuracy in correctly classifying malicious traffic and a 90% chance of detecting malicious traffic using bytes and port numbers, it is obvious that combining the two would provide a very accurate and efficient model for predicting and mitigating Advanced Persistent Threats. These results show greater potentials at mitigating APT than the results from related works in Table 2, leading Chandran et al. [20] work using random forest with 0.14%. The other methods used provided much lower accuracy levels making our approach the most efficient in mitigating APT. The results of the simulations using the four algorithms had acceptable accuracy levels but the increase in the accuracy using the optimized random forest algorithm improves the efficiency in mitigating the threat.

VII. Conclusion

The Advanced Persistent Threats attack increases every year with increasing levels of sophistication. With the inability to detect and prevent these attacks organizations including the government are at a high risk of losing valuable information and services. Having investigated mitigation techniques as highlighted it is evident that there is a need to combine some of the methods highlighted based on their effectiveness. Anomaly detection is the most promising although has some challenges with false positives. Future work in developing a behavioral pattern to reduce the occurrence of false positives will improve the effectiveness in mitigating APT. Utilizing a proposed combination of both static rule-based anomaly detection and machine learning-based techniques showed high accuracy levels in mitigating APT. Being able to detect malicious traffic in 90% of data traffic and a 95% accuracy in detecting malicious activities using machine learning techniques provides an ensemble model for mitigating APT with highly reduced chances for false positives.

References

[1]O. I. Adelaiye, A. Showole and S. A. Faki, ”Evaluating Advanced Persistent Threats Mitigation Effects: A Review,” International Journal of Information Security Science, vol. 7, (4), pp. 159-171, 2018.

[2]R. Brewer, ”Advanced persistent threats: minimising the damage,” Network Security, vol. 2014, (4), pp. 5-9, 2014.

[3]N. Virvilis, B. Vanautgaerden and O. S. Serrano, ”Changing the game: The art of deceiving sophisticated attackers,” in Cyber Conflict (CyCon 2014), 2014 6th International Conference On, 2014, .

[4]M. Ask et al, ”Advanced persistent threat (APT) beyond the hype,” Project Report in IMT4582 Network Security at Gjøvik University College, Springer, 2013.

[5]C. Tankard, ”Advanced persistent threats and how to monitor and deter them,” Network Security, vol. 2011, (8), pp. 16-19, 2011.

[6]M. Nicho and S. Khan, ”Identifying Vulnerabilities of Advanced Persistent Threats: An Organizational Perspective,” International Journal of Information Security and Privacy (IJISP), vol. 8, (1), pp. 1-18, 2014.

[7]I. Jeun, Y. Lee and D. Won, ”A practical study on advanced persistent threats,” Computer Applications for Security, Control and System Engineering, vol. 339, pp. 144-152, 2012.

[8]P. Giura and W. Wang, ”Using large scale distributed computing to unveil advanced persistent threats,” Science J, vol. 1, (3), pp. 93-105, 2012.

[9]S. Singh, Y. Jeong and J. H. Park, ”A survey on cloud computing security: Issues, threats, and solutions,” Journal of Network and Computer Applications, vol. 75, pp. 200-222, 2016.

[10]P. Chen, L. Desmet and C. Huygens, ”A study on advanced persistent threats,” in IFIP International Conference on Communications and Multimedia Security, 2014, .

[11]Y. Su et al, ”A Framework of APT Detection Based on Dynamic Analysis,” 2016.

[12]J. de Vries et al, ”Systems for detecting advanced persistent threats: A development roadmap using intelligent data analysis,” in Cyber Security (CyberSecurity), 2012 International Conference On, 2012, .

[13]G. Brogi and V. V. T. Tong, ”Terminaptor: Highlighting advanced persistent threats through information flow tracking,” in New Technologies, Mobility and Security (NTMS), 2016 8th IFIP International Conference On, 2016, .

[14]P. Lamprakis et al, ”Unsupervised detection of APT C&C channels using web request graphs,” in International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment, 2017, .

[15]I. Friedberg et al, ”Combating advanced persistent threats: From network event correlation to incident detection,” Comput. Secur., vol. 48, pp. 35-57, 2015.

[16]F. Skopik, I. Friedberg and R. Fiedler, ”Dealing with advanced per- sistent threats in smart grid ICT networks,” in Innovative Smart Grid Technologies Conference (ISGT), 2014 IEEE PES, 2014, .

[17]A. Vance, ”Flow based analysis of advanced persistent threats detecting targeted attacks in cloud computing,” in Problems of Infocommuni- cations Science and Technology, 2014 First International Scientific- Practical Conference, 2014, .

[18]Y. Wang et al, ”A network gene-based framework for detecting advanced persistent threats,” in P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), 2014 Ninth International Conference On, 2014, .

[19]I. Ghafir et al, ”Detection of advanced persistent threat using machine-learning correlation analysis,” Future Generation Comput. Syst. vol. 89, pp. 349-359, 2018.

[20]S. Chandran, P. Hrudya and P. Poornachandran, ”An efficient classification model for detecting advanced persistent threat,” in 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2015, .

[21]T. Schindler, ”Anomaly detection in log data using graph databases and machine learning to defend advanced persistent threats,” arXiv Preprint arXiv:1802.00259, 2018.

[22]P. K. Sharma et al, ”DFA-AD: a distributed framework architecture for the detection of advanced persistent threats,” Cluster Computing, vol. 20, (1), pp. 597-609, 2017.

[23]D. Moon et al, ”DTB-IDS: an intrusion detection system based on decision tree using behavior analysis for preventing APT attacks,” The Journal of Supercomputing, vol. 73, (7), pp. 2881-2895, 2017.

[24]M. Ring et al, ”Flow-based benchmark data sets for intrusion detection,” in Proceedings of the in Proceedings of the 16th European Conference on Cyber Warfare and Security (ECCWS). 1em Plus 0.5 Em Minus, 2017, .

[25]T. Rosado and J. Bernardino, ”An overview of openstack architecture,” in Proceedings of the 18th International Database Engineering & Applications Symposium, 2014, .

[26]P. K. Sharma et al, ”DFA-AD: a distributed framework architecture for the detection of advanced persistent threats,” Cluster Computing, vol. 20, (1), pp. 597-609, 2017.