VAWKUM Transactions on Computer Sciences
http://vfast.org/journals/index.php/VTCS@ 2021, ISSN(e):2308-8168, ISSN(p): 2411-6335
Volume 9, Number 1, January-December, 2021 *pp:01-07*

A Survey of Feature Extraction and Feature Selection Techniques Used in Machine Learning-Based Botnet Detection Schemes

Oyelakin A. M.[1], Jimoh R. G. [2]
[1]Department of Computer Science, Al-Hikmah University, Ilorin, Nigeria.
[1]amoyelakin@alhikmah.edu.ng
[2]Department of Computer Science, University of Ilorin, Ilorin, Nigeria
[2]jimoh_rasheed@unilorin.edu.ng

ABSTRACT

Machine learning (ML) techniques are popular for classification of botnets as they have been argued to have improved strengths compared to signature-based approaches. The level of performances of some of these detection schemes have been traced to the relevance of the features used for the classification models. Therefore, extraction and selection of the most discriminative features in the classification of botnets is an important research area. It has equally been found that when a Machine-learning based approach is being used to identify botnets, the dataset chosen has to be real and representative. Extraction and selection of promising features are necessary steps prior to using a Machine Learning-based classification algorithm for identifying botnets. The reason for the pre-processing and feature selection steps in Machine Learning-based model is to be able to remove irrelevant and redundant data in the experimental dataset, minimize computational complexity, and increase both model simplicity as well as accuracy. This paper provided a survey of some feature extraction and feature selection approaches that have been used for ML based botnet detection models. The main purpose of this approach is to provide a better understanding and insights on how improved botnet detection mechanisms can be achieved through enhanced feature extraction and selection methods.

KEYWORDS: Botnet Malware, Machine Learning Algorithms, Netflow Features, Botnet Detection Schemes.

JOURNAL INFO
HISTORY: Received: March 2021 Accepted: July 2021 Published: September 2021

INTRODUCTION

Botnets have been identified as serious threat vectors used for carrying out various attacks by the attackers. Several studies have proposed ML approaches for the identification of botnets in the cyber space. However, botnet malware are becoming more and more sophisticated against detection models as they are coming up with resilient features. Botnets have been identified as the most serious of the malware in the cyber space [1]-[5]. The attacks carried out by botnets include DDOS, Phishing, Identify theft, Spyware, Spoofing, Click fraud and spam attacks [6]. A bot is an autonomous program that is performing some tasks. In a botnet, every malicious bot hides and communicates with its command and control server. A botnet contains a number of these bots and are used for carrying out coordinated attacks in enterprise networks or cyber space.

The number of cyber or network attackers is on the increase and these insecurity issues has led to various intrusion detection approaches [7].With the increase in these botnet attacks, computer networks suffer from attacks that render ICT infrastructure useless [8]. Discovering the most useful and representative features is of great importance for any machine learning problems [9]. Researchers in classification or regression problems have clearly identified feature extraction and feature selection as very key stage in predictive analytics. Feature Selection involves selecting the most relevant attributes while Feature Extraction focuses on combining attributes into a new reduced set of features; Feature Selection focuses on selecting the most relevant attributes or variables from the dataset. Relevant features are extracted and selected using suitable algorithms and will go a long way in achieving improved botnet detection schemes.

There have been several publicly available botnet datasets that have been used for botnet researches in the literature. Some examples are: ISOT botnet dataset [10], ISCX botnet dataset [11], and CTU-13 dataset [12]. However, some of these datasets have been reported to be small or generated using synthetic modelling. The common characteristics of these datasets are they contain a number of features and have to be effectively pre-processed and selected before being fed into a Machine Learning classifier. Out of all these datasets, CTU-13 is currently the largest as it contains real botnet and non-botnet traces [12] as there is a need for improved model for classifying botnets [13]. The CTU-13 dataset was captured at Czech Technical University with a view to advancing researches in botnet detection. The dataset is very large and is contained in thirteen captures, popularly called scenarios. On each of the captures (scenarios) in CTU-13 dataset, a specific malware which used several protocols and performed different malicious actions are included. Based on the characteristics of the dataset, improved feature extraction and selection technique will be required. This paper surveyed different feature extraction and feature selection techniques that have been proposed in Machine-learning based botnet researches.

Importance of Extraction and Selection Features

A feature is the input attribute which serves as a representative information that is extracted from the raw data set. Extraction and selection of good network features which can maximally aid in the detection of malicious traffic [14]. That is why these techniques have been found to be very useful in Machine Learning-based problems across many domains.Authors in [15] claimed that dimensionality reduction is a pre-processing step in ML that is targeted at removing irrelevant so that predictive accuracy can be gained. Similarly,[16] argued that selecting proper features for the classification model is important. However, he further pointed out that there is a trade-off between achieving high detection accuracy.

In most classification tasks, feature extraction is succeeded by Feature selection. The Feature selection algorithm in a Machine Learning-based predictive analysis focuses on selecting promising variables. In Feature Selection, we have the full feature set and then try to build an identified feature set for the problem in the domain that we are investigating [17]. If a researcher uses all the features to build a classification model, this leads to a significant overhead. Similarly, if he uses improper or too few features may cause the accuracy rate to decrease. The choice of which of the approaches is to be used depends on a number of factors. At times, the problem may require that both feature extraction and feature selection methods are used while one of them may be preferred in some other scenarios. However, the emphasis of this work is on works that reported extraction and selection of features in botnet detection researches.

Feature extraction involves the transformation of original dataset to a dataset with a reduced number of attributes, while the feature selection involves the selection of features which contain the most discriminative information. In feature selection, given a feature space, the focus is on how to have an optimal mapping which will serve as the one that does not result into increase in the minimum probability of error. In classification problems, authors in [18] have pointed out that the feature extraction and selection process can be achieved in different ways depending on the goal, and resources. Moreover, [19] mentioned that feature extraction is a very core determinant for the purpose of developing a powerful botnet detection system. This is because to be able to design an effective botnet detection model, it is required to identify which discriminating features to get.

Machine Learning-based Classification Problems

Common problem in ML is defeminising a representative set of features that allow us to construct a classification model for a task [20]. The classification focuses on learning from a training dataset *S* in which each instance is represented by feature-values with a target class , and to predict class labels of the instances outside *S.* The increase in the use of internet platforms for various activities has equally attracted people with malicious intent to launch attacks in the cyber space through the use of malware such as viruses, trojan horses, botnets and many others [21]. Before a Machine Learning researcher begins building a model using a dataset, it is important that a preliminary investigation of the data be carried out so as to better understand its specific characteristics.

The Machine-learning based detection models make use of datasets while making botnet prediction. These datasets are expected to be real and comprehensive [22]. While designing and developing detection models, Feature Extraction and Selection techniques have been found to have great impact. This is because discovering the most useful and a representative feature is of great importance for any machine learning problems [23]. In most learning-based tasks across different research domains, data can be represented by a fixed number of features which can be binary, categorical or continuous. It is then left with the researcher to find a good data representation which most times, is very domain-specific. In the literature, researchers have been employing machine learning methods using different feature sets as a prevalent method [24]. It is believed that improvement in feature extraction techniques will promote the correct classification of modern botnet malware.

Machine Learning Algorithms and Botnet Detection Models

Machine Learning techniques have become prominent for use in detecting botnets in the internet space [23]. However, due to the increasing evolution and sophistication of these botnets, improved models are required from time to time. A Machine Learning algorithm can be supervised, unsupervised or semi-supervised. Authors in [24] [25] asserted that high dimensionality of data is a fundamental problem in science applications that use data as a result of growth in learning complexity and this is applicable in botnet detection because modern botnets are becoming more sophisticated in their approaches to attacks.

The best way to approach any classification problem in ML is to start by doing a proper analysis of the dataset. The main focus of a feature extraction technique is to reduce the number of features in a dataset by creating new features from the existing ones and then the throw the original features out. It has to be pointed out that in any ML classification problem, we do have a pairs of observations that are taken as (x_i,y_i) which are drawn from distributions such as (blood status, cancer) and similar classification problems in different domains.

While using Machine Learning approach for the detection of botnets, it is important to make sure that the true representations of features are used.Apart from the Traditional Machine Learning algorithms such as Support Vector Algorithms, Naive Bayes, Logistic Regression, Random Forests, Deep Learning Algorithms such as Deep Belief Network (DBN), Recurrent Neural Network (RNN), and Convolution Neural Network (CNN) are fast being used in botnet detection researches. It has been

argued that Deep Learning extend the state of the art in feature extraction as Deep Learning classifiers have the ability to scale when the datasets are large enough [26].These Deep Learning Algorithms allow automatic feature extraction or construction from the domain datasets. That is, Deep Learning algorithms have come with the power of automatically extracting features in your data unlike the traditional Machine Learning algorithms. Figure 1 shows the fundamental principle upon which a Machine Learning model operates.

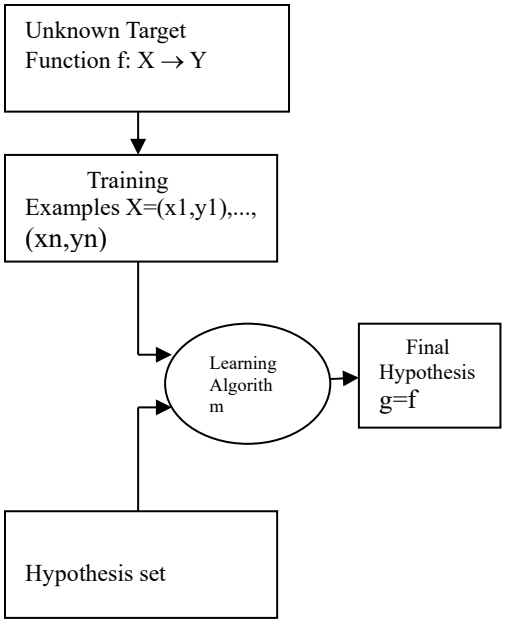

Figure 1: Basic set up of ML Problem [36]

METHODOLOGY

The methodology adopted in this study is a survey of related literature on the subject-matter. The survey focuses on feature extraction and feature selection approaches that have been proposed in past botnet classification studies. Recent studies in botnet domain in respect of feature extraction and feature selection methods were emphasised in the work. Emphasis was on relevant papers in notable research outlets from 2013 to 2019 were consulted. A feature extraction method explains the process through which targeted features are obtained out of the network traces while the feature selection describes the process through which a sub set of features are identified from the whole.

3.0 FEATURE EXTRACTION AND SELECTION TECHNIQUES IN BOTNET DETECTION MODELS

Researchers in [27] proposed a feature extraction method for Peer-to-Peer Botnet Detection using Graphic Symmetry Concept. The authors used this approach due to the various kinds of features required to identify a Peer-to-Peer (P2P) botnet characteristics. Botnets are known to be used for various attacks such as DDoS attacks, phishing attacks and much more. The paper extracted data packet size and corresponding periods in the flow according to the concept of graphical symmetry. Combined with the entropy of the flow information and session features, frequency domain features can be sorted to obtain features with better correlations, solving the problem of the multiple types of features required for bot detection. The authors in [28] designed an approach to extract different predefined types of network events from network flow data. The proposed technique is based on analyzing the computational properties of event types as stipulated by their attributes in a specific descriptive language. The corresponding events are then extracted with a higher recall compared to a relevant event extraction portion of the intrusion detection system during production called Camnep.

In the study by [29], a number of features were extracted for the purpose of identifying evidence of botnet. Features were registered using nfdump. The researchers in [29] used net streaming features that have the following attributes: Ts (stream start time), Te (stream ends time), Td (stream duration), Sa (source IP address), Da (destination IP address), Sp (source port), Dp (destination port), Pr (protocol), Ra / flg (Flg Flags), Ipkt (input packets), ibyt (input bytes). The authors used the filter selection method, to identify discriminatory traits. This technique is applied to measure the subset of features with the highest

predictive power. Every filter based feature is not based on any algorithm and can be used in any field. The authors used CfsSubsetEval as a filter method to select the most relevant traits in the experiment [29]. The researchers argued in [30] that because HTTP connections created as command and control (C&C) channels in Botnets have specific implementations that differ from the patterns used by legitimate HTTP accesses, a bot proof can be created. The authors in [8] used connection-based and HTTP-based features that are able to distinguish between C&C bot channels and benign traffic. The authors in [8] also claimed that the method achieved more than 99% accuracy with zero false positives, based on the proposed approach.

Authors in [16] developed a botnet detection scheme by analyzing their network flows on the Internet. The system classified similar network flow traffic into groups, and then extracted the behavior patterns of each group for machine learning classifier. The system focused on analyzing Peer-to-Peer botnets characteristics the flow traffic features. Similarly, authors in [31] used a netflow-based technique for the detection of botnets with the use of selected machine learning algorithms to software defined networks without the reading of packet payload. The study used network flows as input and process these flows so as to extract a statistical feature set to be used for classification. Network flow stream was used for the feature extraction. The feature set that was extracted was from selected trace and relevant historical flows. In [9][23], the researchers carried out a study that investigated the presence of botnets in CTU-13 dataset and focus of the extraction of network flow summary in the dataset. The features were extracted using a Deep Learning architecture. The chosen Deep Learning architecture was able to do feature extraction automatically which is different from the techniques in Shallow Machine Learning Algorithms.

Authors in [11] designed a technique that generated feature set for botnet detection that is based on connectivity among botnets at their phase of C and C server. The authors aimed at maximizing the detection rate of these botnets. A Genetic Algorithm was used to select the set of features that gives the highest detection rate. In [9] , the writers analyzed the most discriminating features for the purpose of building an efficient and effective botnet detection system. The authors argued that there is not enough research performed to explore the effect of the selection of feature set in botnet detection. Authors in [11] reviewed studies on netflow-based features employed in the existing botnet detection studies and then evaluated their relative effectiveness. The researchers were able to show that the effectiveness of different combination of features in terms of providing more detection coverage has not been fully studied [11].

Writers in [32] proposed a study that is based on the analysis of network flow community behaviour. Authors focused on community behaviour of the attributes if the network features generated with a view to identifying the evasion techniques used by Peer-to-Peer botnets. Authors in [33] investigated how the packet header features extraction extracted through the use of Netmate and Tranalyzer affect the detection accuracy of the selected Machine Learning algorithms. Experimentations carried out on two public botnet datasets show that the feature extraction approaches affect the performances of the selected Machine Learning algorithms: Support Vector Machine, Logistic Regression and Neural Network. The work concluded that the extracted features from the botnet dataset have great impact on the performances of the botnet detection algorithms. The researchers in [34] emphasized the need to have an improved botnet detection models because new variant of botnets are exhibiting evasion techniques.

The authors in [35] - [38] expressed traffic features as flows using the Softflowd tool. In this case, the features are derived from the package header information alone. The authors also claimed in [38] that the features can be used to classify encrypted traffic as well. Also, researchers in [14] performed a performance comparison of three different feature selection algorithms in a machine learning-based classifier. These feature selection techniques include correlation-based feature selection; Consistency-based subgroup assessment and principal component analysis. The authors then used the features identified in some machine learning algorithms. Experimental comparisons were made on three different machine learning techniques - decision trees, Naïve Bayes classifier, and Bayesian network classifier. Classifiers were evaluated to detect the traffic of peer-to-peer (P2P) botnets. In the study, the authors emphasized that a number of training data samples and a good "feature set" are prerequisites for building effective classification models using machine learning algorithms.

DISCUSSION

Some of the literature surveyed are very current and discussed the evolving directions in Machine learning-based botnet detection researches. Machine Learning-based botnet detection mechanisms that were found in literature were examined and reported. The studies focus on feature extraction and selection techniques used in some of these Machine-Learning based detection models. The reviewed papers emphasized the need for efficient and effective pre-processing of botnet datasets for the purpose of having a more accurate. To have a generally acceptable datasets for Intrusion Detection System, the datasets used for botnet detection are required to be comprehensive enough in order for them to aid in the detection accuracy. Discovering the most useful and representative features in such datasets is of great importance for any machine learning-based model [9]. Some of the studies reviewed equally pointed out that there is a trade-off between achieving high detection accuracy and spending huge computation time on constructing a large classification model. It was equally observed that Net flow features were the central focus of most of the works reviewed.

For this reason, researchers that use all features in a botnet dataset to build a Machine classification model can have significant overhead. As pointed out by [36], a Machine Learning problem follows the steps in figure 1. Some of the works

reviewed also pointed out that using improper or too few features may cause the detection accuracy rate of the model to decrease. The review equally pointed out that CTU-13 dataset is getting more attention from recent botnet researchers as the dataset is argued to be more comprehensive and realistic in nature when compared to others such as ISCX, ISOT and so on. Studies such as [9] [23][37] emphasized the imperative of using ML techniques for botnet detection just like it was emphasized in this survey work.

CONCLUSION AND FUTURE WORK

The work discussed the feature extraction and feature selection approaches reported in literature in respect of Machine Learning-based botnet detection models. The study identified that most authors emphasised how good feature extraction and selection for different botnet datasets can aid in achieving improved detection of the malware in the cyber space. During the review of literature, it was revealed that CTU-13 dataset is becoming very popular for botnet researches as it has been identified to contain real-life traces and found to be very comprehensive. Specifically, the future work will focus on achieving improved selection of the most discriminative features in the evolving botnet (CTU-13) dataset. That is, in future, our interest is to design improved feature selection approaches that can be used for selecting discriminative features in a benchmark dataset named CTU-13. Then, such features will be used to achieve improved botnet detection mechanisms in more innovativeway. Based on the different protocols in the dataset, we intend focusinon features that can promote the detection of some resilient botnets that use P2P and DNS-based protocols. It is believed that this approach will go a long way in helping the security research communities to achieve more successes in the fight against botnet malware.

ACKNOWLEDGMENT

The first author was supported in part by 2019/2020 Jim Ovia-NCS PhD Computer Science Scholarship at University of Ilorin, Ilorin, Nigeria. Thanks are also due to the anonymous reviewers who helped in improving the manuscript through constructive comments.

REFERENCES

[1] P. B. and V. Yegneswaran, “P. Barford and V. Yegneswaran .” An Inside Look at Botnets”, to appear in Series: Advances in Information Security. Springer, 2006,” *Springer*, 2006.

[2] D. Julian B. Grizzard, V. Sharma, Nunnery C., “Julian B. Grizzard, V. Sharma, Nunnery C., Kang B.B. & Dagon D. ‘Peer-to-Peer Botnets: Overview and Case Study’, Proceedings of the First Conference on First work on Hot Topics in Understanding Botnets, 2007,retrieved from https://pdfs.semanticscholar.org.”

[3] J. Liu, Y. Xiao, K. Ghaboosi, D. Hongmei, “J. Liu, Y. Xiao, K. Ghaboosi, D. Hongmei, and J. Zhang. Botnet. ‘Classification, Attacks, Detection, Tracing and Preventing Measures’. Journal on Wireless Communication and Networking, 2009,” p. 2009.

[4] M. Muhammad, N. Manjinder, “‘A Survey on Botnet Architectures, Detection and Defences’, International Journal of Network Security, 0(0), PP.1-19,2013,” *Int. J. Netw. Secur.*, pp. 1–19, 2013.

[5] W. Ping, W. Lei, “Analysis of Peer-to-Peer Botnet Attacks and Defences, Department of Electrical Engineering and Computer Science, 2010,” *Dep. Electr. Eng. Comput. Sci.*, 2010.

[6] D. Deeper, “Digging Deeper – An In-Depth Analysis of a Fast Flux Network, Akamai White Paper,2006,” 2006.

[7] D.-h Lee,., D.-y Kim,., “Multi-Stage Intrusion Detection System Using Hidden Markov Model Algorithm. International Conference on Information Science and Security , 72-7,2008,” 2008.

[8] M. D. Santana, “What we Learn from Learning-Understanding Capabilities and Limitations of Machine Learning in botnet attacks,2018, retrieved from https://arxiv.org/abs/1805,on 26th August 2018,” p. 2018.

[9] A. P. and T. A. T., “‘Effective Feature Selection for Botnet Detection Based on Network Flow Analysis.’ International Conference Automatics and Informatics’,2017,” *Int. Conf. Autom. Informatics*, 2017.

[10] A. Alenazi A., I. Traore , K. Ganame, “‘Holistic Model for HTTP Botnet Detection Based on DNS Traffic Analysis’. In: Traore I., Woungang I., Awad A. (eds) Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments. ISDDC 2017. Lecture Notes in Computer Science, vol 10618,” *Springer, Cham*, 2017.

[11] E.B. Beigi, H.H. Jazi, N. Stakhanova, “‘Towards effective feature selection in machine learning-based botnet detection approaches’.2014 IEEE Conference on Communications and Network Security, CNS 2014, 247–255 (2014), https://doi.org/10.1109/CNS.2014.6997492,” *CNS*, pp. 247–255, 2014, doi: 10.1109.

[12] A. Z. . Sebastian Garcia, Martin Grill, “Sebastian Garcia, Martin Grill, Jan Stiborek and Alejandro Zunino . ‘An empirical comparison of botnet detection method’, Computers and Security Journal, Elsevier, 45, 100,2014 123. http://dx.doi.org/10.1016/j.cose.2014.05.011,” *Elsevier*, 2014, doi: 10.1016.

[13] B. S. Lagraa, J. Francois, A. Lahmadi, “S. Lagraa, J. Francois, A. Lahmadi, M. Miner. BotGM : Unsupervised Graph Mining to Detect Botnets in Traffic Flows, HAL Id : hal-01636480 , 2017,” 2017.

[14] P. Narang, J.M. Reddy, “‘Feature selection for detection of peer-to-peer botnet traffic’ Compute 2013 - 6th ACM India

Computing Convention: Next Generation Computing Paradigms and Technologies. (2013), https://doi.org/10.1145/2522548.2523133," 2013, doi: 10.1145/2522548.2523133.

[15] K. Samina, K. Tehmina & N. Shaomila." A Survey of Feature Selection and Feature Extraction Techniques in Machine Learning", 2014 Science and Information Conference," *Sci. Inf. Conf.*, 2014.

[16] C. Hung, "C. Hung, & H. Sun."A Botnet Detection System Based on Machine-Learning using Flow-Based Features", SECURWARE 2018: The Twelfth International Conference on Emerging Security Information, Systems and Technologies, 122–127.," *Twelfth Int. Conf. Emerg. Secur. Information, Syst. Technol.*, pp. 122–127, 2018.

[17] M. Ved., "M. Ved." Feature Selection and Feature Extraction in Machine Learning: An Overview", retrieved from https://medium.com/@mehulved1503/feature-selection-and-feature-extraction-in-machine-learning-an-overview-57891c595e96."

[18] A. Jović, A., ""A review of feature selection methods with applications". 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2015 - Proceedings, 1200–1205(2015). https://doi.org/10.1109/MIPRO.2015.7160," *MIPRO*, 2015, doi: 10.1109.

[19] J. Jianguo, B. Qi, S. Zhixin, Y. Wang, "' Botnet detection method analysis on the effect of feature extraction', Proceedings - 15th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, 10th IEEE International Conference on Big Data Science and Engineerin," *Int. Symp. Parallel Distrib. Process. with Appl. IEEE*, pp. 1882–1888, 2016.

[20] M. A. Hall., "Correlation-based Feature Selection for Machine Learning, a PhD Thesis at University of Waikato,(1999)," 1999.

[21] K. Shanthi., "Categories of Botnets, World Academy of Science, Engineering and Technology, International Journal of Computer and Systems Engineering 8(9), 1689–1692, 2014," *Int. J. Comput. Syst. Eng.*, vol. 8, no. 9, pp. 1689–1692, 2014.

[22] M Marek, "Network Intrusion Detection: Half a Kingdom for a Good Dataset,https://pdfs.semanticscholar.org/b39e/0f1568d8668d00e4a8bfe1494b5a32a17e17.pdf?_ga=2.237473350.756880770.1576358584-422052986.1572640169," 2015.

[23] A. Pektas and T. Acarman T. 'Effective Feature Selection for Botnet Detection Based on Network Flow Analysis.' International Conference Automatics and Informatics'," 2017.

[24] T. A. Pektaş, & T. Acarman, "' Botnet detection based on network flow summary and deep learning', International Journal of Botnet Detection in Software Defined Networks. International Journal of Security and Its Applications, 11(11), 1–12. (2017)https://doi.org/10.14257/ijsia.2017.11," vol. 11, no. 11, pp. 1–15, 2018.

[25] R. Bellman, "Dynamic Programming. Princeton, NJ: Princeton University Press, 1957," 1957.

[26] T. Epelbaum, "Deep Learning: Technical Introduction, 2017 September," *Tech. Introd.*, 2017.

[27] Z. Yang, "" A Feature Extraction Method for P2P Botnet Detection Using Graphic Symmetry Concept", .Symmetry, 11(3), 326, https://doi.org/10.3390/sym11030326, 2019," vol. 11, no. 3, 2019.

[28] G. Sourek and F. Zeleny, "" Efficient Extraction of Network Event Types from NetFlows, Security and Communication Networks, 2019,https://doi.org/10.1155/2019/8954914," *Secur. Commun. Networks,* 2019.

[29] P. A. (2018). L. Mathur, M. Raheja, "Botnet Detection via mining of network traffic flow, Procedia Computer Science 132:1668-1677, DOI: 10.1016/j.procs.2018.05.137," 2018.

[30] P. A. A. Resende, "HTTP and contact-based features for Botnet detection. Security and Privacy, 1(5), e41 (2018). https://doi.org/10.1002/spy2.41," *Secur. Priv.*, vol. 1, no. 5, 2018.

[31] F. Tariq, &S. Baig, "Machine Learning Based Computers,CONIELECOMP(2017). https://doi.org/10.1109/CONIELECOMP.2017.7891834," *CONIELECOMP(2017).*, 2017.

[32] D. Zhuang,, & J.M. Chang. 'Detecting Peer-to-Peer Botnets through Community Behavior Analysis' 2017 IEEE Conference on Dependable and Secure Computing, 493–500. (2017), http://doi.org/10.1109/DESEC.2017.8073832," pp. 493–500, 2017.

[33] J. Jianguo, B. Qi, S. Zhixin, Y. Wang, "J. Jianguo, B. Qi, S. Zhixin, Y. Wang, & B. Lv. ' Botnet detection method analysis on the effect of feature extraction', Proceedings - 15th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, 10th IEEE Internation," *IEEE*, 2016.

[34] R. G. Jimoh., "A. M. Oyelakin & R. G. Jimoh. 'A Review on the Identification Techniques for Detection-Evasive Botnet Malware', in the proceedings of International Conference of Nigeria Computer Society, Gombe, Nigeria, July 2019," *Int. Conf. Niger. Comput. Soc. Gombe, Niger.*, 2019.

[35] F. Haddadi, D. Runkel, "'On botnet behaviour analysis using GP and C4.5.GECCO', 2014 - Companion Publication of the 2014 Genetic and Evolutionary Computation Conference, 2014, 1253–1260. https://doi.org/10.1145/2598394.2605435," *Evol. Comput.*, 2014.

[36] M. M.-I. & H. T. L. Y.S. Abu-Mostafa, "Y.S. Abu-Mostafa, M. Magdon-Ismal & H.T. Lin. Learning from data. AML Book, 2012," *AML B.*, 2012.

[37] M. Pederson., "M. Stevanovic & J.M. Pederson. "On the use of Machine learning for identifying botnet network traffic:, Journal of Cyber Security, Vol. 4, 1–32. 2016, doi: 10.13052/jcsm2245-1439.421," *J. Cyber Secur.*, vol. 4, pp. 2245–1439, 2016.

[38] F. Haddadi, D. Runkel, "] F. Haddadi, D. Runkel, A. Nur Zincir- Heywood, Malcolm I. Heywood. 'On botnet behaviour analysis using GP and C4.5', Proceedings of the 2014 conference companion on Genetic and evolutionary computation companion - GECCO Comp '14, 2014," 2014.