International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT206444

Efficient Cognitive Fog Computing for Classification of Network Cyberattacks Using Machine Learning

Prof. A. V. Deorankar, Shiwani S. Thakare
Computer Science and Engineering, Government College of Engineering, Amravati, Maharashtra, India

Article Info
Volume 6, Issue 4
Page Number : 176-184
Publication Issue :
July-August-2020

Article History
Accepted : 20 July 2020
Published : 25 July 2020

ABSTRACT

IoT is the network which connects and communicates with billions of devices through the internet and due to the massive use of IoT devices, the shared data between the devices or over the network is not confidential because of increasing growth of cyberattacks. The network traffic via loT systems is growing widely and introducing new cybersecurity challenges since these loT devices are connected to sensors that are directly connected to large-scale cloud servers. In order to reduce these cyberattacks, the developers need to raise new techniques for detecting infected loT devices. In this work, to control over this cyberattacks, the fog layer is introduced, to maintain the security of data on a cloud. Also the working of fog layer and different anomaly detection techniques to prevent the cyberattacks has been studied. The proposed AD-IoT can significantly detect malicious behavior using anomalies based on machine learning classification before distributing on a cloud layer. This work discusses the role of machine learning techniques for identifying the type of Cyberattacks. There are two ML techniques i.e. RF and MLP evaluated on the USNW-NB15 dataset. The accuracy and false alarm rate of the techniques are assessed, and the results revealed the superiority of the RF compared with MLP. The Accuracy measures by classifiers are 98 and 53 of RF and MLP respectively, which shows a huge difference and prove the RF as most efficient algorithm with binary classification as well as multi- classification.

Keywords : Cybersecurity, Fog layer, IoT, Anomaly detection.

I. INTRODUCTION

IoT (Internet of Things) is an increasingly popular term and also known as umbrella term which covers technologies and smart devices that both have Internet capabilities. Smart devices are easier to use and more comfortable and thus gain more popularity to make our life easier. On the other side, the increased deployment of smart devices brings an increase in potential security risks. Due to massive use of IoT devices the network complexity is increased, and it becoming harder to manage future network due to network cyber attacks. Cyber attacks are used to obtain unauthorized access to the IoT devices without the knowledge of either the eligible user or administrator and due to this the fog layer is

introduced. Fog layer is used to reduce the energy consumption, latency and storage. The very most important goal of fog layer is to improve security, efficiency and reduce the amount of data that needs to be sending to the cloud for processing, analysis and storage. But mostly it is done for security and efficiency reason. Fog computing is a concept, which analyze the origin of data from the outer edges, fog layer will analyze that where data is created and where it will be store either in the cloud or in a customer's data center. The primary goal of fog layer is to improve the efficiency and to reduce the redundancies related with data being transferred to cloud for its process and to be stored, which will maximize the security as compared to cloud computing. The data will not directly send to the cloud layer, since it creates a high latency network connection between devices and analytics endpoints as well as the larger amount of bandwidth as compare to fog layer. Most importantly in some scenarios there is no bandwidth connection to send data, because the data is being processed from where it is created. To manage the network and a security vulnerabilities, an IDS is used, which monitors the data traffic in order to identify and protect the systems information. The operations of IDS are divided into three stages. The first stage is monitoring stage, the second stage is the analysis stage and the final stage is detection stage. The architecture is based on the advantage of fog computing to reduce the latency between cloud and IoT sensor. It comprises of three layers that include application layer, fog layer and IoT sensor layer. The Fog layer is a major component of the architecture, which ensures processing and aggregation of the data. The AD-IoT system is designed to monitor all IoT traffic in a distributed fog layer and alert the administrator or the service provider.

1. Things Layer/ End Devices:

End devices are Smart Devices which are small and consist of sensors, controllers, Actuators i.e. memory constrained. Things layer of IoT comprises of electronics devices. The smart devices can be phones or tablets, micro controller units and single-board computers. The devices which are connected are the real endpoint of IoT. These devices include subsystem, sensors, embedded device, mobile device, etc. The main capabilities of a typical IOT device are: Devices should be able to sense and record data, to perform light computing and finally and being able to connect to a network and communicate the data.

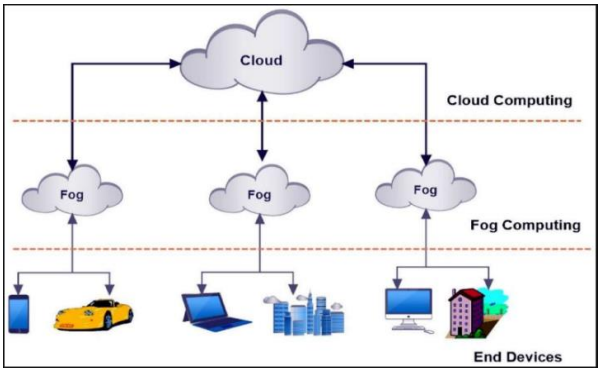

Figure 1: Architecture: level of communication

2. Fog Layer:

Fog layer is used to reduce the energy consumption, latency and storage. The very most important goal of fog layer is to improve security, efficiency and reduce the amount of data that needs to be sending to the cloud for processing, analysis and storage. But mostly it is done for security and efficiency reason. Fog computing is a concept, which analyze the origin of data from the outer edges, fog layer will analyze that where data is created and where it will be store either in the cloud or in a customer's data center. The primary goal of fog layer is to improve the efficiency and to reduce the redundancies related with data being transferred to cloud for its process and to be stored, which will maximize the security as compared to cloud computing. Fog layer consist of certain IoT services, like prediction, monitoring, planning, inferring, diagnosis, maintenance i.e. pre-processing, which is closer to

edges so that it enables a faster local automation and decision making.

3. Cloud Layer:

Cloud computing consist of large no of computing task which share high speed resources to a large scale computing centers and virtualization technology, which will help to reduce hardware maintenance costs and computing time. Cloud Computing is a model which provide remote access through internet to share media in the form of services. Cloud host servers are there, which consist of all sensor data, where cloud host server store data and process data for analysis and decision making.

II. RELATED WORK

This section presents previous studies that relates to this research of anomaly detection techniques based on machine learning algorithm in IoT network. A distributed deep learning has driven fog-to-things computing attack detection scheme using NSL-KDD dataset. To compare model with shallow algorithms, metrics such as accuracy, DR, and ROC curve have been used for system evaluation, and accuracies over varied worker nodes are considered for scalability measure [2]. Although the results are not yet good to be adopted in any commercial product for signature-based IDS, the approach still has significant potential land advantages for further development [3].

According to the prior studies the traditional Intrusion Detection System (IDS) method consist of Host Based intrusion detection system(HIDS) and Network based intrusion detection system(NIDS) or hybrid (IDS), which can detect cyberattacks in different ways, where as traditional IDS are designed to detect intrusion activities on a single or whole network traffic. The types of IDS work accordingly, that is, the first type is Host Based IDS, which install software like anti-virus and detect the suspicious activities of the network traffic, by scanning and analyzing the different activities such as system call, application logs, file system etc. And these activities are not significant with some IoT devices and thus this method get fails due to limited functionality and resources [4]- [8].The second type is network based IDS, which monitors entire network traffic and detect known and unknown attacks unlike (HIDS), based on hybrid method which has both anomaly based technique and signature based technique [6], [9],[10].Signature based method consumes more power and fails to detect attack, it detects only the records which is stored in the database[18].The anomaly based NIDS method is more efficient for monitoring the network traffic as well as detecting new attacks. Therefore, anomaly detection technique promises to detect attacks by using NIDS method.

Recently the Random Forest (RF) algorithm is used to detect any malicious behavior by using the UNSW-NB15 dataset, and AD-IoT detection method, where the binary classification is used for classifying the behavior of the packet whether it is normal or malicious. But, it is only restricted for behavior of packets. There is no knowledge of exact attack takes place on any packet [11].

The AD-IoT technique is not yet been used with multi-layer Perceptron algorithm to detect any malicious behavior as well as to specify the name of attack, so this will be the new research according to the prior researches[20].

III. METHODOLOGY

This section provides the development of proposed system. This chapter describes the detailed information about the development of the proposed approach. These are further discussed in brief.

In this Anomaly detection model, the flow of detection starts from the very first step i.e. IoT layer which is also known as Things layer. This layer of IoT comprises of devices, sensors and controllers. Connected devices which enable the IoT environment. These devices include mobile devices such as smart phones or tablets, micro controller units and single-board computers. The connected devices are the real endpoint for IoT. Things layer of IoT comprises of electronics devices. These electronics devices are the small, memory-constrained and consist of sensors and actuators. These devices include subsystem, sensors, embedded device, mobile device, etc. The next layer is fog layer, Fog layer of certain IoT services, like prediction, monitoring, planning, inferring, diagnosis, maintenance i.e. pre-processing, which is closer to edges so that it enables a faster local automation and decision making. And the last layer is cloud layer, Cloud computing allows a large number of computing tasks to share high-speed hardware resources through the establishment of large-scale computing center and virtualization technology, which can effectively reduce computing and hardware maintenance costs. The layer focuses more on the application of high latency data with a large number of data types and complex computational model.

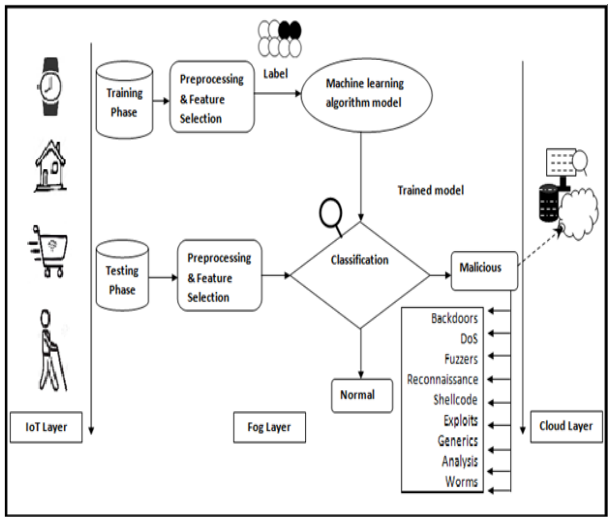

Figure 2: AD-IoT detection system model

Cloud computing is actually a model for the availability and use of Information and Communication Technologies, which enables remote access via the Internet to a range of shared computing media in the form of services. All the sensor data is stored on cloud hosted servers, which store and process data for analysis and decision making.

We propose a detection method system called AD-IoT for detecting cyberattacks at fog nodes in a network as shown in fig. 2. The framework of this method relies on different machines learning algorithm to enhance the efficiency of Ad-IoT network. We proposed this method works to monitor the network traffic that passes through each fog node, as fog nodes are nearest to IoT sensors, rather than detection on the huge amount of the cloud storage to identify among normal and abnormal behaviors. After detecting attacks in the fog level, it should alert the security cloud services to inform them to analyze and update their system. The machine learning algorithm which has been used in this work is random forest (RF) and multi-layer perceptron (MLP) with multi classification. According to the result, it has been proven that the accuracy, processing time, performance, etc., of Random forest algorithm is better than that of multi-layer perceptron. The better result we have achieved from the Random Forest.

IV. EVALUATION AND EXPERIMENTAL RESULTS

This section presents the analysis and evaluation of our proposed framework, Ad-IoT, on different parameters which based on the UNSW-NB 15 dataset. This study evaluated machine learning algorithm to identify cyber attack and the category of cyberattacks and network traffic from malicious activities to apply the final model on fog nodes to the AD-IoT approach in future work.

i. Dataset Description

The UNSW-NB 15 dataset was discovered in 2015 which was used to address the issues of modern normal and malicious network cyber attack in network traffic. The UNSW-NB 15 dataset have a benchmark over the older dataset. The older dataset are KDDCUP 99 and NSLKDD, these datasets are widely use before UNSW-NB 15 dataset to evaluate NIDS performance. But they get failed to provide the realistic output performance due to lots of reasons. Whereas the UNSW-NB 15 dataset is more capable of providing the realistic output performance. The dataset consist of nine types of attacks, namely, Fuzzers, Analysis, Backdoors, DoS, Shellcode, Exploits, Generic, Reconnaissance and Worms and 49 features with the class label. The number of records in the training set is 175,341 records and the testing set is 82,332 records from the different types, attack and normal.

TABLE I. Attack Category in UNSW-NB 15

Type of Attacks	No. Records
Normal	2,218,761
Fuzzers	24,246
Analysis	2,677
Backdoors	2,329
DoS	16,353
Exploits	44,525
Generic	215,418
Reconnaissance	13,987
Shellcode	1,511
Worms	174

In this paper, firstly we have taken complete 82,332 records of dataset to train the model and for testing purpose 175341 records are used by using RF and MLP classifiers. In the second step we are doing the classification according to the attack category. Since we are performing multi classification there are total nine types of attacks and we are analyzing each of them and the output is shown in table V given below for each category of attack.

ii. Evaluation Metrics

This paper measures the model to present the results by using confusion matrix and other metrics such as Precision, Recall, F1 score, Hit Rate, Miss Rate, Positive Predictive Value, Negative Predictive Value, False Discovery rate and false omission rate**.** Firstly the efficiency of proposed model is count by AD-IoT for detecting cyber attack and is shown by counting the instance in actual normal records by correctly and incorrectly detecting the instance.

TABLE II. Confusion Matrix

		Predicted	
		Negative	Positive
Actual	Negative	TN	FP
	Positive	FN	TP

A confusion matrix contains information about actual and predicted classifications done by a classification system. Performance of such systems is commonly evaluated using the data in the matrix. The above table II shows the confusion matrix. According to the values from confusion matrix the parameter such as TN, TP, FN, FP are calculated and the results are shown in the figure 3 and 4 of RF and MLP respectively given below.

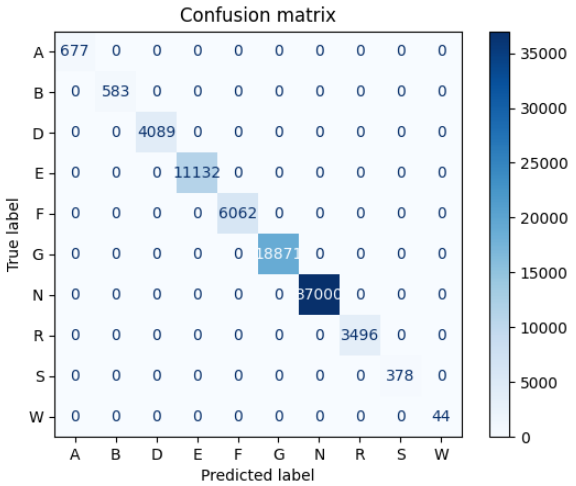

Fig 3: Confusion Matrix for RF test

Figure 3 represents the confusion matrix of RF Classifier, in which x-axis represents Predicted label and y-axis represents True Label. Where A, B, D, E, F, G, N, R, S, W represents cyber attacks i.e. Analysis = 667, Backdoors = 583, DoS = 4089, Exploit = 11132, Fuzzers = 6062, Generics = 18873, Normal = 87000, Reconnaissance = 3496, Shellcode = 978 and Worms = 44.

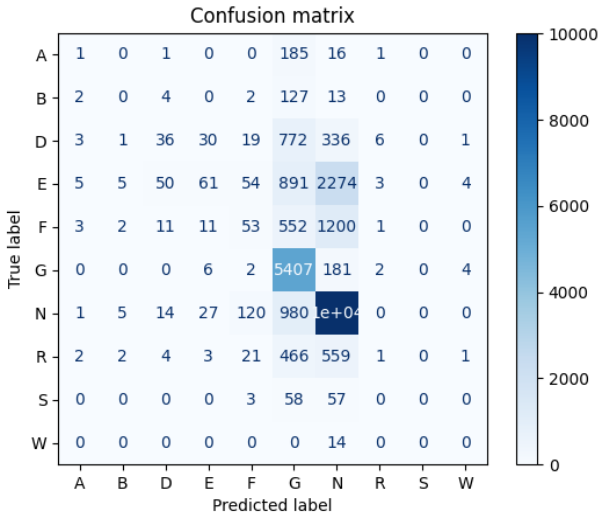

Fig 4: Confusion Matrix for MLP test

Figure 4 represents the confusion matrix of MLP Classifier, in which x-axis represents Predicted label and y-axis represents True Label. Where A, B, D, E, F, G, N, R, S, W represents cyber attacks i.e. Analysis = 1, Backdoors = 0, DoS = 36, Exploit = 61, Fuzzers = 53, Generics = 5407, Normal = 10024, Reconnaissance = 1, Shellcode = 0 and Worms = 0.

According to this TN, TP, FN, FP the statistical parameter are calculated as shown in table III and IV of RF and MLP respectively, the statistical parameter are Precision, Recall, F1 score, Hit Rate, Miss Rate, Positive Predictive Value, Negative Predictive Value, False Discovery rate and False omission rate.

TABLE III. Parameters Evaluation Using RF

Class Name	Precision	Recall	F1-Score	Hit Rate	Miss Rate	Positive Predictive Value	Negative Predictive Value	False Discovery Rate	False omission Rate
Analysis	76	60	67	60.25	39.75	75.54	99	24.45	00.45
Backdoors	99	45	62	45.07	54.92	99.24	99	0.75	00.54
DoS	91	100	95	99.55	44.03	91.26	99	0.87	00.03
Exploits	98	100	99	99.62	37.43	98.33	99	0.16	00.08
Fuzzers	98	98	98	98.47	15.23	98.25	99	0.17	00.01
Generic	100	99	99	99.25	75	99.70	99	0.02	00.02
Normal	100	100	100	99.96	39.28	99.70	99	0.02	00.00
Reconnaissance	99	99	99	98.63	13.63	99.06	99	0.09	00.08
Shellcode	99	88	93	87.81	12.18	98.61	99	0.01	00.07
Worms	96	76	85	76.15	23.84	96.11	99	0.03	00.00

TABLE IV. Parameters Evaluation Using MLP

Class Name	Precision	Recall	F1-Score	Hit Rate	Miss Rate	Positive Predictive Value	Negative Predictive Value	False Discovery Rate	False omission Rate
Analysis	4	1	1	0	99	0.40	98	95	0.11
Backdoors	2	0	0	0	99	0.16	99	98	0.00
DoS	26	3	6	3	96	2.58	93	74	0.68
Exploits	45	2	5	2.4	97	4.46	81	55	1.887
Fuzzers	27	3	5	2.7	97	2.66	89	73	1.01
Generic	53	98	69	98	0.1	5.33	99	46	0.00
Normal	56	95	70	94	0.54	5.55	96	44	0.38
Reconnaissance	8	0	0	0	99.76	0.77	94	92	0.59
Shellcode	0	0	0	0	100	-	99	-	0.00
Worms	0	0	0	0	100	-	99	-	0.00

TABLE V. COMPARATIVE ANALYSIS OF RF AND MLP

PARAMETER NAME	AVERAGE PERFORMANCE	
	RF	MLP
ACCURACY	0.983	0.536
PRECISION	0.956	0.442
RECALL	0.865	0.404
F1-SCORE	0.897	0.312
HIT RATE	86.47	20.01

MISS RATE	35.52	78.84
Positive Predictive Value	95.85	2.191
Negative Predictive Value	99	94.7
False Discovery Rate	2.65	57.7
False omission Rate	0.128	0.465
MCC	0.9795	0.40736
KAPPA SCORE	0.9794	0.36177

Table V shows the comparison between both the classifiers i.e. RF and MLP. The results of RF are almost perfect and the results of MLP are below the expectations. The Accuracy of RF is 98% and MLP consist only 53% which is very poor percentage. Matthews Correlation Coefficient which is used to Calculate Average Accuracy, and for RF the MCC is 97.95% and for MLP it is only 36.17%. The accuracy and false alarm rate of the techniques are assessed, and the results revealed the superiority of the RF compared with MLP, which shows a huge difference and prove the RF as most efficient algorithm with binary classification as well as with multi-classification.

V. CONCLUSION AND FUTURE WORK

Now a day, we are facing a cyberattacks via IoT devices in network traffic and thus introduced an approach based on NIDS called AD-IoT system to detect various loT attacks in a distributed fog layer instead of a cloud layer. To reduce this the AD-IoT technique is used based on machine learning through the evaluation of the UNSW-NB 15 dataset to make data more securable. The proposed AD-IoT can significantly detect malicious behavior using anomalies based on machine learning classification before distributing on a cloud layer. This work discusses the role of machine learning techniques for identifying the type of Cyberattacks. There are two ML techniques i.e. RF and MLP evaluated on the USNW-NB15 dataset. The accuracy and false alarm rate of the techniques are assessed, and the results revealed the superiority of the RF compared with MLP. The Accuracy measures by classifiers are 98 and 53 of RF and MLP respectively, which shows a huge difference and prove the RF as most efficient algorithm with binary classification as well as multi-classification.

VI. FUTURE SCOPE

In the near future, we will try to implement an application base system with real time data which will evaluate the detection accuracy and computational efficiency. There is still room for improvement in terms of detection accuracy and computational efficiency by adding more features or using more efficient algorithms. One possible direction is to use a more efficient learning algorithm.

VII. REFERENCES

[1]. Abeshu and N. Chilarnkurti, "Deep learning: the frontier for distributed attack detection in fog-to-things computing," IEEE CommunicationsMagazine, vol. 56, no. 2, pp. 169-175, 2018.

[2]. T. A. Tang, L. Mhamdi, D. McLernon, S. A. R. Zaidi, and M. Ghogho, "Deep learning approach for network intrusion detection in software defined networking," in Wireless Networks and Mobile Communications (WINCOM), 2016 International Conference on. IEEE, 2016, pp. 258-263.

[3]. B. Zarpelao, R. S. Miani, C. T. Kawakani, and S. C. de Alvarenga, "A survey of intrusion detection in internet of things;' Journal of Networkand Computer Applications, vol. 84, pp. 25-37, 2017.

[4]. M. M. Hossain, M. Fotouhi, and R. Hasan, "Towards an analysis of security issues,

challenges, and open problems in the internet of things," in Services (SERVICES), 2015 IEEE World Congress on. IEEE, 2015, pp. 21-28.

[5]. H. Habibzadeh, T. Soyata, B. Kantarci, A. Boukerche, and C. Kaptan, "Sensing, communication and security planes: A new challenge for a smart city system design r," Computer Networks, vol. 144, pp. 163- 200, 2018.

[6]. N. Moustafa and J. Slay, "Unsw-nb15: a comprehensive data set for network intrusion detection systems (unsw-nb15 network data set)," in Military Communications and Information Systems Conference (MilCIS),2015. IEEE, 2015, pp. 1-6.

[7]. N. Koroniotis, N. Moustafa, E. Sitnikova, and J. Slay, "Towards developing network forensic mechanism for botnet activities in the iot based on machine learning techniques," in Mobile Networks and Management:9th International Conference, MONAMI 2017, Melbourne, Australia,December 13-15, 2017, Proceedings, vol. 235. Springer, 2018, pp. 30--44.

[8]. M. Nobakht, V. Sivaraman, and R. Boreli, "A host-based intrusion detection and mitigation framework for smart home iot using openflow," in Availability, Reliability and Security (ARES), 2016 11th InternationalConference on. IEEE, 2016, pp. 147-156.

[9]. H. Summerville, K. M. Zach, and Y. Chen, "Ultra-lightweight deep packet anomaly detection for internet of things devices;' in Computing and Communications Conference (IPCCC), 2015 IEEE 34th International Performance. IEEE, 2015, pp. 1-8

[10]. S. Garg, K. Kaur, N. Kumar, S. Batra, and M. S. Obaidat, "Hyclass: Hybrid classification model for anomaly detection in cloud environment," in 2018 IEEE International Conference on Communications (ICC). IEEE, 2018, pp. 1-7.

[11]. M. Ahmed, A. N. Mahmood, and J. Hu, "A survey of network anomaly detection techniques," Journal of Network and Computer Applications, vol. 60, pp. 19-31, 2016. [l] J. Howell. Number of connected iot devices will surge to 125 billion by 2030, ihs markit says - ihs technology. [Online]. Available: https://technology.ihs.com/596542, last accessed: 11/07/2018.

[12]. Borgia, "The internet of things vision: Key features, applications and open issues," Computer Communications, vol. 54, pp. 1-31, 2014.

[13]. Restuccia, S. D'Oro, and T. Melodia, "Securing the internet of things: New perspectives and research challenges," IEEE Internet of ThingsJournal, vol. 1, no. 1, pp. 1-14, 2018.

[14]. J. A. Stankovic, "Research directions for the internet of things," IEEE Internet of Things Journal, vol. 1, no. 1, pp. 3-9, 2014.

[15]. M. Antonakakis, T. April, M. Bailey, M. Bernhard, E. Bursztein, J. Cochran, Z. Durumeric, J. A. Halderman, L. Invernizzi, M. Kallitsis et al., "Understanding the mirai botnet;' in USENIX Security Symposium, 2017, pp. 1092-1110.

[16]. J. Santos, P. Leroux, T. Wauters, B. Volckaert, and F. D. Turck, "Anomaly detection for smart city applications over 5g low power wide area networks," in NOMS 2018 - 2018 IEEE/IFIP Network Operationsand Management Symposium, 2018, PP-1.

[17]. Yousefpour, G. Ishigaki, and J. P. Jue, "Fog computing: Towards minimizing delay in the internet of things," in Edge Computing (EDGE),2017 IEEE International Conference on. IEEE, 2017, pp. 17-2.

[18]. R. Roman, J. Zhou, and J. Lopez, "On the features and challenges of security and privacy in distributed internet of things," ComputerNetworks, vol. 57, no. 10, pp. 2266-2279, 2013.

[19]. S. Prabavathy, K. Sundarakantham, and S. M. Shalinie, "Design of cognitive fog computing for intrusion detection in internet of things," Journal

of Communications and Networks, vol. 20, no. 3, pp. 291-298, 2018.

[20].D. Oh, D. Kim, and W.W. Ro, "A malicious pattern detection engine for N. Moustafa and J. Slay, "The evaluation of network anomaly detection systems: Statistical analysis of the unsw-nb15 data set and the comparison with the kdd99 data set," Information Security Journal: A GlobalPerspective, vol. 25, no. 1-3, pp. 18-31, 2016.

[21].M. M. Rathore, A. Paul, A. Ahmad, N. Chilarnkurti, W.-H. Hong, and H. Seo, "Real-time secure communication for smart city in high-speed big data environment," Future Generation Computer Systems, vol. 83, pp. 638-652, 2018.

[22].J. Zhang, M. Zulkernine, and A. Haque, "Random-forests-based network intrusion detection systems," IEEE Transactions on Systems, Man, andCybernetics, Part C (Applications and Reviews), vol. 38, no. 5, pp. 649-659, 2008.

[23].K. Angrishi, "Turning internet of things (iot) into internet of vulnerabilities (iov): lot botnets," arXiv preprint arXiv:1702.03681, 2017.

[24].Alqazzaz, I. Alrashdi, E. Aloufi, M. Zohdy, and H. Ming, "Secsps: A secure and privacy-preserving framework for smart parking systems," Journal of Information Security, vol. 9, no. 04, pp. 299-314, 2018.

[25].A. V. Deorankar and S. S. Thakare, "Classification of Network Cyberattacks for Efficient Cognitive Fog Computing," 2020 Recent Innovations in Wireless Network Security Volume 2 Issue 1, HBRP Publication Page 1-7 2020, DOI: http://doi.org/10.5281/zenodo.3767885

[26].A. V. Deorankar and S. S. Thakare, "Survey on Anomaly Detection of (IoT)- Internet of Things Cyberattacks Using Machine Learning," 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2020, pp. 115-117, doi: 10.1109/ICCMC48092.2020.ICCMC-00023.

Cite this article as :

Prof. A. V. Deorankar, Shiwani S. Thakare, "Efficient Cognitive Fog Computing for Classification of Network Cyberattacks Using Machine Learning", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 6 Issue 4, pp. 176-184, July-August 2020. Available at doi : https://doi.org/10.32628/CSEIT206444
Journal URL : http://ijsrcseit.com/CSEIT206444