Towards Security Enhancement for NFV-Based IoT Networks Using Machine Learning

Sandeep N. Gite[1](✉) and Smita L. Kasar[2]

[1] Maharashtra Institute of Technology, Aurangabad, Maharashtra, India
sandeepgite37@gmail.com
[2] Maharashtra Institute of Technology, Aurangabad, Maharashtra, India

Abstract. IoT networks are increasing day-to-day life and due to its growing use their sizes are increasing continuously. As IoT network contains different devices supporting different platforms these networks are vulnerable to different types of attacks. There is a need for an intrusion detection system that can recognize these attacks and carry out appropriate defenses against them in order to secure the network. Several other systems with individual classifiers have been tried, but they are insufficient to recognize attacks correctly and carry out defenses against them. This system introduces an ensemble classifier that integrates different classifiers and optimizes their parameters with Harris Hyena optimization. This optimization will improve classification accuracy and the rate of detection will be fast which helps to execute proper countermeasures against a particular attack In this work, these challenges are considered, and planning to introduce an efficient artificial intelligence (AI) approach to address the three key IoT security-related issues. Along with the AI approach, the NFV (Network function virtualization) will be integrated to develop a generalized Intrusion Detection System (IDS). Malware and virus present in the network are detected by this architecture. Lastly, the NFV surveillance zone patch structure-based IoT network's scalability is assessed. This procedure aims to create an algorithm that can manage all IoT devices that are currently operating. The basic objective of this method is to develop a hybrid machine-learning algorithm that can support most, if not all, of the IoT devices, now being utilized for diverse purposes. This model will forecast malware assaults for NFV so that suitable defenses can be put in place. Data Controller distributes updates through its distance-bound service as soon as the malware virus is discovered in a particular base station database. An NFV service-based framework that is distance-bound allows for effective update distribution. The system handles a wide variety of adware and viruses that can be found. It provides NFV with a hybrid machine-learning algorithm and patching system to prevent malware spread in IoT networks.

Keywords: Intrusion Detection System (IDS) · Network Function Virtualization (NFV) · BoT-IoT · SEIR Model · Denial of Service (DoS)

B. Raj et al. (Eds.): ICETE 2023, AER 223, pp. 361–369, 2023.
https://doi.org/10.2991/978-94-6463-252-1_40

1 Introduction

Our everyday lives are embracing IoT devices as the new standard. With the help of IoT technology, it is possible to create a smarter environment that, among other things, may help people save time, energy, and money [11]. The cornerstone of cloud computing technology is virtualization, which shares a pool of reconfigurable resources to offer on-demand access to a variety of programs and services. Due to the growing size of Internet of Things (IoT) networks and the corresponding rise in data consumption, a Network Function Virtualization (NFV) system offers the best security solution in terms of overcoming complexity and creating efficient attack prediction models to process and patch malware attacks[12] [13][1]. Despite the fact that IoT devices enhance consumer quality of life, they are prone to security breaches and are readily exploitable if security countermeasures are not implemented. For instance, researchers found that 178 million IoT devices, such as webcams, medical devices, routers, and more, are accessible to attackers in ten US locations via the public internet. [14] Operating systems built on Linux are used by the majority of IoT devices. As a result, the developers of malware attempt to create the source code while taking into account the flaws found in the operating systems and apps used by IoT devices. [3].

Because of the growing size of IoT networks and the associated rise in data consumption, a Network Function Virtualization (NFV) system provides the best security solution in terms of overcoming complexity and developing efficient attack prediction models to process and patch malware attacks. IoT vulnerability-based risks can be divided into four categories because of the extensive connectivity between IoT devices and the Internet:

(1) Denial of Service (DoS)
(2) Malware attacks
(3) Data breaches
(4) Weakening Perimeters

2 Related Work

Many companies and organizations have deployed software and services onto virtualized cloud platforms as a result of the rapid evolution of cloud computing in recent years. Static and dynamic approaches are the two main categories of malware detection techniques [16]. In the first, characteristics are extracted from the target executable files using static analysis. Byte sequences, strings, and disassembled instructions make up the majority of these characteristics [4].

Network administration and management are made simple using Software-Defined Networking (SDN), which separates the control plane from the data plane. Exposing programmable interfaces, it makes the creation and deployment of network apps easier. SDN is susceptible to DoS saturation assaults, however [17] [5]. Network function virtualization (NFV) has a broad range of applications today, from mobile core networks to IP node implementations (such as the future Internet architecture) [18]. Network function virtualization (NFV) enables network operations to be carried out in virtual machines (VMs) within a cloud architecture as opposed to specific hardware [19].

Because smart devices only have a small quantity of memory, storage, and processing power, it is difficult to detect attacks. It is challenging to produce effective malware signatures for IoT devices due to their heterogeneous processor design [21]. Unfortunately, organizations primarily employ manual processes to deal with malware-related incidents, so they are unable to stop current attacks or prospective ones from happening in the future. Because the control and data planes are separated in SDN, the network may be programmed and applications and network services can be supplied by abstracting away the underlying architecture [7]. A large number of devices may now be infected by a single DDoS attack. For additional information, see the following: by saturating the network with requests, zombies attack the victim and prevent real users from utilizing services [22].

To take control of a collection of compromised computers that are connected to the same network, attackers deploy malicious software, also referred to as the master DDoS [23][8]. Due to the expanding possibilities, network operators and service providers are racing to implement network-slicing features. With the help of network slicing, numerous logical networks can be easily created on top of a physical network infrastructure that is shared by all users. SDN-based architecture is more susceptible to malicious attacks than monolithic core architecture as a result of network function virtualization's increased availability of network infiltration sites for attackers [9]. The table below lists existing systems related to our work (Table 1).

After the study of all these related systems, we can conclude that there is a need for an optimized machine learning system that can integrate different classifiers together to get an ensemble classifier and solve the classification problem and countermeasure execution problem of existing systems. It can also improve the detection rate through optimization methods. From this research gap need, we formulate the objectives of our work as follows:

1. To effectively perform network virtualization in data controller using optimized ensemble classifiers.
2. To develop an ensemble model by hybridizing the classifiers effectively.
3. To optimize the ensemble classifier efficiently using Harris Hyena optimization techniques.
4. To effectively evaluate the proposed model using various evaluation metrics.

3 Proposed Solution

The research's primary objective is to detect malware origins in order to improve security for network function virtualization. The system model of the network is shown in Fig. 1. The system model consists of multiple IoT devices and the devices are connected to the base stations. The data controller controls the communication between all the base stations. Virtualization will be provided to genuine devices and illegal access will be prevented in this research. The data from the data controller will be aggregated and preprocessed to reduce the complexity of the detection.

The output of the pre-processing step will be supplied into the ensemble machine learning model, which is built by mixing various classifiers including the light GBM, neural network, decision tree, and K-nearest neighbor. These classifiers can be hybridized to effectively distinguish between typical and abnormal users. The Harris hyena optimization is developed from the standard hybridization of the spotted hyena optimization

Table 1. An Inventory of Current Systems Relevant to Our System

Sr. No	Author	Methods	Advantage	Disadvantage
1.	Nadra Guizani and Arif Ghafoor [1].	Recurrent Neural Network Model short-term memory (RNN-LSTM).	Improved the security and performance of the IoT device network. Increased malware attacks.	Low accuracy for the classification of the malware group.
2.	Tianliang Lu *et al.* [2]	A combined Deep Belief Network-Gate Recurrent Unit Model.	Higher detection accuracy, and better detection effect on obfuscated malware.	The pace of learning is excessive. The algorithm could not reach convergence. A slow gradient decline may occur.
3.	Rajasekhar Chagantia *et al.* [3]	A neural network with bidirectional-gated recurrent units and convolutions (Bi-GRU-CNN).	Reduced the high dimensionality data.	Some malware families in the dataset are difficult for us to categorize properly.
4.	Donghai Tian *et al.* [4]	*Miami-Dade County Health Department* (MDCHD)	Improved the performance of our detection mechanism and minimal performance cost.	The checking procedure is relatively slow.
5.	Samer Y Khamaiseh et al. [5]	Machine learning-based saturation attack detection system.	Lower rate. Reduced the occurrence of table miss packets.	Reduced detection performance.
6.	Mikhail Zolotukhin and Timo H¨amalainen [7]	Software-defined networking and network function virtualization.	Improved network security in 5G networks.	The average packet size is lower and the system effectively drops slow connections.

(*continued*)

and Harris hawk optimization will tune the weights and bias of the classifiers and provide an optimized output. The Hybrid Machine Learning Algorithm will be used to successfully optimize the various parameters that each classifier uses. The hybridized ensemble classifier predicts the normal and abnormal users. If it is a normal user then the virtualization will be performed and if it is identified as an abnormal user then the devices will furthermore seek virtualization from the data controllers. The research will

Table 1. (*continued*)

Sr. No	Author	Methods	Advantage	Disadvantage
7.	Ahamed aljuhani [8]	DDoS (Distributed denial of service) attack mitigation technique.	The environment offers, flexibility, scalability, and cost reduction.	Higher performance rate and lower delays. The web server response is very slow.
8.	Muder Almiani et al. [10]	Kalman backpropagation neural network-based DDoS (Distributed denial of service).	High detection performance and high predictive accuracy.	Not only did the requirement for a lot of training material decrease.

be carried out using the software Python and the efficiency will be proved by measuring accuracy, sensitivity, specificity, and time delay.

Network Function Virtualization (NFV) will be used for experimentation where an IoT network with all accessories will be simulated. A model will be utilized in conjunction with botnet attacks to test the system and identify various malware attack types. A group of Internet-connected devices that are joined to a botnet each operates one or more bots.

Any assault that utilizes a botnet, which is a group of linked devices and bots that perform a single job for distribution and scalability, is referred to as a botnet attack. Massive cybercrimes like DDoS and severe scraping are carried out by cybercriminals via botnet attacks. The data controller gets the data from all base stations, this data will be aggregated and preprocessed for better classification of data.

This Preprocessed data will be forwarded to the ensemble classifier which consists of light GBM, KNN, DT, NN, etc. classifiers. Harris Hyena optimization will be applied to the parameters of a classifier to improve the performance of classification and increase the detection rate. A source of the attack is identified malware containing data that will not be forwarded by the data controller and that device is called a seek device or

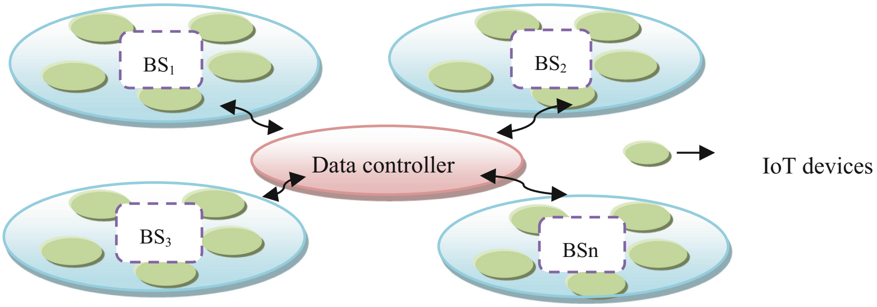

Fig. 1. Security Enhancement Model for IoT Network

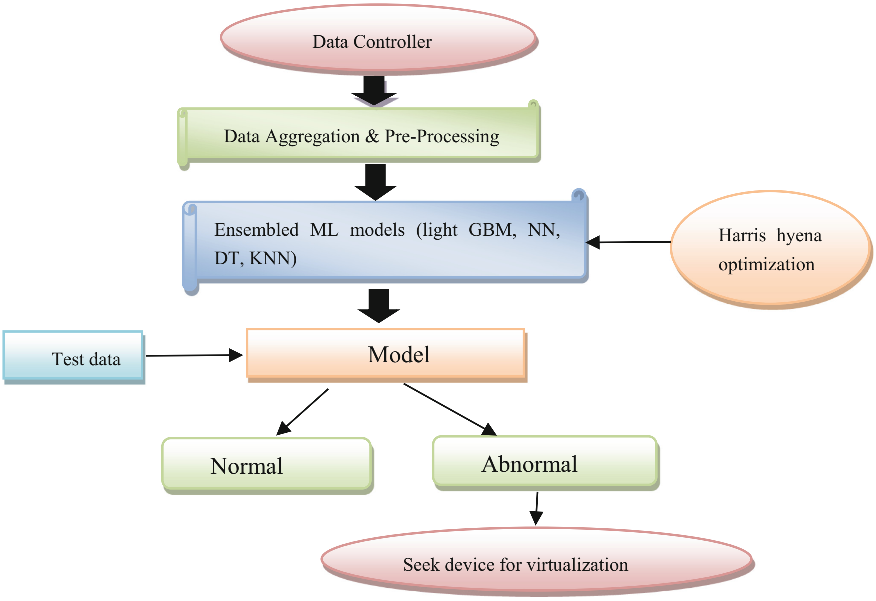
Data Controller
Data Aggregation & Pre-Processing
Ensembled ML models (light GBM, NN, DT, KNN)
Harris hyena optimization
Test data
Model
Normal
Abnormal
Seek device for virtualization

Fig. 2. Proposed System Architecture

treated as abnormal for virtualization. If data classified by the data controller is normal corresponding data is forwarded by and data transmission is completed as shown in Fig. 2.

While going through related technologies to reach the objectives we can face challenges as:

1. IoT technology integration poses a number of security risks since it gives malevolent people a way to access, modify, and use sensitive personal data.
2. IoT devices also employ a range of CPU architectures and operating systems. Therefore, using traditional malware detection and classification methods to identify and categorize IoT malware is difficult.
3. Due to a number of security issues, the NFV is susceptible to some cyber assaults.
4. The difficulties in implementing NFV in the real world for both performance optimization and diagnostic reasons.

4 Conclusion

The research addresses the need for developing a hybrid machine learning-based malware & virus detection method. The problem in the research is regarding the detection of the attack in IoT networks and enhancing the detection accuracy. Further, categorization of the attack in a large IoT network is a challenging task in the detection process. Hence, a new model is required for the effective identification of malware and virus attacks to execute appropriate countermeasures and provide a patching system.

This method's main objective is to develop a hybrid machine-learning algorithm that can support most of IoT devices. This model will forecast malware assaults for NFV so that suitable defences can be put in place. The virtualized system releases updates via the service as soon as the malware virus is found. Update distribution is made possible via a distance-bound NFV service-based infrastructure. The system handles a wide variety of adware and viruses that can be found.

References

1. Guizani, Nadra, and Arif Ghafoor, "A network function virtualization system for detecting malware in large IoT based networks," IEEE Journal on Selected Areas in Communications, vol. 38, no. 6, pp. 1218-1228, 2020.
2. Lu, Tianliang, Yanhui Du, Li Ouyang, Qiuyu Chen, and Xirui Wang, "Android malware detection based on a hybrid deep learning model," Security and Communication Networks, vol. 2020, pp. 1-11, 2020.
3. Chaganti, Rajasekhar, Vinayakumar Ravi, and Tuan D. Pham, "Deep learning based cross architecture internet of things malware detection and classification," Computers & Security, vol. 120, pp. 102779, 2022.
4. Tian, Donghai, Qianjin Ying, Xiaoqi Jia, Rui Ma, Changzhen Hu, and Wenmao Liu, "MD-CHD: A novel malware detection method in cloud using hardware trace and deep learning," Computer Networks, vol. 198, pp. 108394, 2021.
5. Khamaiseh, Samer Y, Izzat Alsmadi, and Abdullah Al-Alaj, "Deceiving machine learning-based saturation attack detection systems in sdn," In 2020 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN), IEEE, pp. 44–50, 2020.
6. Firoozjaei, Mahdi Daghmehchi, Jaehoon Paul Jeong, Hoon Ko, and Hyoungshick Kim, "Security challenges with network functions virtualization," Future Generation Computer Systems, vol. 67, pp. 315-324, 2017.
7. Zolotukhin, Mikhail, and Timo Hämäläinen, "On artificial intelligent malware tolerant networking for IoT," In 2018 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN), IEEE, pp. 1–6, 2018.
8. Aljuhani Ahamed, "Machine learning approaches for combating distributed denial of service attacks in modern networking environments," IEEE Access, vol. 9, pp. 42236-42264, 2021.
9. Thantharate, Anurag, Rahul Paropkari, Vijay Walunj, Cory Beard, and Poonam Kankariya, "Secure5G: A deep learning framework towards a secure network slicing in 5G and beyond," In 2020 10th annual computing and communication workshop and conference (CCWC), IEEE, pp. 0852–0857, 2020.
10. Almiani, Muder, Alia AbuGhazleh, Yaser Jararweh, and Abdul Razaque, "DDoS detection in 5G-enabled IoT networks using deep Kalman backpropagation neural network," International Journal of Machine Learning and Cybernetics, vol. 12, pp. 3337–3349, 2021.
11. M. S. Mahdavinejad, M. Rezvan, M. Barekatain, P. Adibi, P. Barnaghi, and A. P. Sheth, "Machine learning for internet of things data analysis: A survey," Digital Communications and Networks, vol. 4, no. 3, pp. 161–175, 2018.
12. M. Ojo, D. Adami, and S. Giordano, "A SDN-IoT architecture with NFV implementation," in 2016 IEEE Globecom Workshops (GC Wkshps), IEEE, pp. 1–6, 2016.
13. I. Farris, T. Taleb, Y. Khettab, and J. Song, "A survey on emerging SDN and NFV security mechanisms for IoT systems," IEEE Communications Surveys & Tutorials, vol. 21, no. 1, pp. 812–837, 2018.
14. Charlie Osborne F, "Researchers discover over 170 million exposed IoT devices in major US cities — ZDNet," 2017.
15. Flexera, "RightScale 2019 state of the cloud report, https://resources.flexera.com/web/media/documents/rightscale-2019-state-of-the-cloud-reportfrom-flexera.pdf,2019.

16. Anusha Damodaran, Fabio Di Troia, Corrado Aaron Visaggio, Thomas Austin, and Mark Stamp, “A comparison of static, dynamic, and hybrid analysis for malware detection,” J. Comput. Virol. Hacking Techn, vol. 13, pp. 1–12, 2017.
17. R. Swami, M. Dave, and V. Ranga, “Software-defined networking-based DDoS defense mechanisms,” ACM Computing Surveys (CSUR), vol. 52, no. 2, pp. 1–36, 2019.
18. “ETSI. Network Functions Virtualisation; Introductory White Paper. Technical report, SDN and OpenFlow World Congress,” 2012.
19. “ETSI. Network Functions Virtualisation (NFV); Infrastructure Overview. Technical report, ETSI GS NFV-INF,” 2015
20. M. Aiash, G. Mapp, and O. Gemikonakli, “Secure Live Virtual Machines Migration: Issues and Solutions,” In 27th International Conference on Advanced Information Networking and Applications Workshops, vol. 0, pages 160–165, 2014.
21. M. Alhanahnah, Q. Lin, Q. Yan, N. Zhang, and Z. Chen, Efficient signature generation for classifying cross-architecture iot malware,” In 2018 IEEE Conference on Communications and Network Security (CNS), pages 1–9, May 2018.
22. G. Somani, M. S. Gaur, D. Sanghi, M. Conti, M. Rajarajan, and R. Buyya, “Combating DDoS attacks in the cloud: Requirements, trends, and future directions,” IEEE Cloud Comput, vol. 4, no. 1, pp. 22-32, 2017,
23. R. Vishwakarma and A. K. Jain, “A honeypot with machine learning based detection framework for defending IoT based botnet DDoS attacks,” in Proc. 3rd Int. Conf. Trends Electron. Infor-mat. (ICOEI), pp. 1019–1024, 2019.
24. E. Gelenbe and Y. Yin, “Deep learning with dense random neural networks,” in International Conference on Man–Machine Interactions. Springer, 2017, pp. 3–18.
25. R. Nix and J. Zhang, “Classification of android apps and malware using deep neural networks,” in 2017 International joint conference on neural networks (IJCNN). IEEE, 2017, pp. 1871–1878.
26. Y. Meidan, M. Bohadana, A. Shabtai, J. D. Guarnizo, M. Ochoa, N. O. Tippenhauer, and Y. Elovici, “Profiliot: a machine learning approach for IoT device identification based on networktraffic analysis,” in Proceedings of the symposium on applied computing. ACM, 2017, pp. 506–509.
27. J. Canedo and A. Skjellum, “Using machine learning to secure IoT systems,” in 2016 14th Annual Conference on Privacy, Security and Trust (PST). IEEE, 2016, pp. 219–222.
28. M. Nauman, T. A. Tanveer, S. Khan, and T. A. Syed, “Deep neural architectures for large scale android malware analysis,” Cluster Computing, vol. 21, no. 1, pp. 569–588, 2018.
29. F. Martinelli, F. Marulli, and F. Mercaldo, “Evaluating convolutional neural network for effective mobile malware detection,” Procedia computer science, vol. 112, pp. 2372–2381, 2017.
30. X. Hu, X. Li, E. C.-H. Ngai, V. C. Leung, and P. Kruchten, “Multidimensional context-aware social network architecture for mobile crowd sensing,” IEEE Communications Magazine, vol. 52, no. 6, pp. 78–87, 2014.