

Activation functions for deep learning: an application for rare attack detection in wireless local area network (WLAN)

Viet-Thang Vu¹, Thanh Quyen Bui Thi ², Hong-Seng Gan³, Viet-Vu Vu⁴, Do Manh Quang¹,
Vu Thanh Duc⁵, Dinh-Lam Pham^{4,6}

¹Faculty of Information Technology, Hanoi University of Industry, Vietnam

²Institute of Information Technology, Vietnam Academy of Science and Technology, Vietnam

³School of AI and Advanced Computing, XJTLU Entrepreneur College (Taicang), Xi'an Jiaotong - Liverpool
University, Suzhou, Jiangsu Province, P. R. China 215400

⁴VNU Information Technology Institute, Vietnam National University, Hanoi, Vietnam

⁵ Information Technology Center, Ha Noi University of Industry, Vietnam

⁶Contents Convergence Software Research Institute, Kyonggi University, South Korea
vuvietthang@hau.edu.vn

Abstract - As we know, there are now more and more cyberattacks that have a serious impact on individuals, businesses, and governments. Attacks are increasingly sophisticated and modern, requiring competent entities to study and devise solutions to quickly detect and minimize user risks. In fact, attacks come in many different ways and many different environments. In this paper, we do an empirical study on the deep learning model using some kinds of activation functions applied for detecting attacks in the wireless local area network (WLAN) data. We also point out some discussions about activation functions for future research.

Keywords – WLAN, Deep learning, Intrusion Detection System, Rare Attack, activation function.

I. INTRODUCTION

WLAN wireless networks, also known as Wi-Fi local area networks, are the most common type of connection commonly used at the household or work scale and in public places. This is a network using wireless technologies with a coverage range of less than 200m. Due to the popularity as well as the open space of the wireless network (easy scale), it is also the type of network that is vulnerable to intrusion/attack due to many subjective users or lack of knowledge about setting up encryption, and security for connection.

In Vietnam, according to statistics by the end of August 2020, the Center for Cyber Surveillance and Response- SOC (Viettel Company) has detected a total of more than 3 million cyber-attack alerts on financial systems, banks, and information technology networks in some provinces and cities across the country. The three forms of cyber-attacks that financial systems and banks

face most are web mining (accounting for 77.58%), malicious code (12.05%), and depletion (3.92%).

Within the scope of this paper, we focus on and pay attention to the types of attacks in WLAN networks, especially rare attacks because there are not too many studies in this field, most of the topics related to wired networks are mainly, and if there is, the results are not too high. Moreover, the rate of cyber-attacks has increased steadily over the years, which proves that there has never been an algorithm that is optimal and that every system needs to be improved and updated regularly. Rare attacks are understood to be previously unseen forms of attack or very small numbers [1][2][3]. These types of attacks are often very dangerous and have serious consequences because the identification system does not have any information or characteristics on them. In terms of data, this type of attack also has a very small number compared to another popular type.

According to the studies of Kolias et al. [4], for WLAN networks, there are many different types of attacks but can be grouped into 3 main groups: Flooding, Injection, and Impersonation (Table 1).

- Flooding (DDOS): a form of denial-of-service attack. In attacks of this type, the attacker tries to fill the bandwidth of the WLAN to the maximum to prevent traffic from ordinary users, so the WLAN will be unavailable and unable to serve the client.
- Injection: an attacker uses malicious code to take control of the access point and change the configuration of the access point and collect data transmitted via WLAN.
- Impersonation: These attacks are also known as spoofing attacks. The attacker masquerades as one

of the network nodes to get information sent to and from the network nodes.

Table 1. Common types of attacks in WLAN

Group of attacks	Types of attack
Group “Flooding”	Deauthentication, CTS, RTS, Disassociation, Beacon, Deauthentication broadcast, Disassociation broadcast, Probe Request, Block Acknowledge, Authentication Request, Fake Power Saving, Probe Response
Group “Injection”	ARP-Injection, ChopChop, Fragmentation
Group “Impersonation”	Honeypot, Evil Twin, Caffé Latte, Hirte

Table 1 presents some types of attacks in WLAN. In recent years, to prevent network attacks, we can use some machine learning methods to build a model applied for intrusion detection system [9-20].

In this paper, we do an empirical study on the deep learning model using some kinds of activation functions applied for detecting attacks in the wireless local area network (WLAN) data. We also point out some discussions about activation functions for future research.

This paper is organized as follow. The section II presents the activation functions. The section III shows the experimentation results. The section IV concludes the paper.

II. ACTIVATION FUNCTIONS

The aim of activation function in a neural network is to define how the weighted sum of the input is mapped into an output from a node or nodes in a layer of the network. The choice of activation function has a crucial impact on the performance of the system using a neural network model, and different activation functions may be used in different parts of the model. We will brief survey some main kinds of activation function as follows.

The first one is ReLU function. This function proposed by Vinod Nair and Geoffrey E. Hinton in [5] and widely use in deep learning models. The properties of ReLU function are continuous, not-bounded and not zero-centered. When $x = 0$, the left-hand derivative of

ReLU is zero while right hand derivative is 1. The ReLU function is presented in the equation 1.

$$f(x) = \max(0, x) = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \quad (1)$$

The second one is Leaky ReLU (LReLU) [6], The properties of the LReLU are continuous and not-bounded. The computation using LReLU is cheap and having a zero-centered activation function. The LReLU function allows a small part of negative units instead of pushing them to zero like the ReLU function does. The equation 2 shows the definition of the LReLU.

$$f(x) = \max(ax, x) = \begin{cases} x & \text{if } x \geq 0 \\ ax & \text{if } x < 0 \end{cases} \quad (2)$$

The third one is ELU (Exponential Linear Unit) function [7]. In contrast to ReLUs, ELUs have negative values which allows them to push mean unit activations closer to zero like batch normalization but with lower computational complexity. ELUs performs not only to faster learning, but also to significantly better than ReLUs and LReLU on networks having more than 5 layers. The ELU function is expressed in the equation 3.

$$f(x) = \begin{cases} x & \text{if } x \geq 0 \\ \alpha(e^x - 1) & \text{if } x < 0 \end{cases} \quad (3)$$

The last one is SegRELU proposed by Gan et al in [8]. SegRELU has some properties: (1) parameter-free setting to ease users’ implementation, (2) light architecture to encourage reproducibility in future deep learning models, and (3) robust weight update at both the positive and negative part of the activation functions to preserve effective feature learning during training process. The SegRELU is tested for medical image segmentation data. The definition of SegRELU is illustrated in the equation 4.

$$f(x) = \begin{cases} \frac{x}{1+|x|} & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases} \quad (4)$$

As mentioned in recent literature, The activation functions play an important role in machine learning models. For each type of data, the result highly depends on the choice of activation functions, the next section will present an empirical study on 4 kind of activation function applied for intrusion detection data in WLAN.

III. EXPERIMENTATION

To conduct the experiment, we used a part of the AWID dataset (Aegean WiFi Intrusion Dataset) that was actually collected from attacks in the wireless network, denoted O-AWID. AWID consists of 156 attributes, we select 13 features according to the study of Liu et al.

[21]. The characteristic of these data set is presented in the table 2.

To evaluate the performance, the Accuracy and F1-score is used in our experimentation. We use a CNN model with 4 kinds of activation as presented in the section III. The detail architecture of the CNN model and the code can be found in [22].

Table 3-6 shows the results obtained by the system. In each table, we present the results produced when using 4 kind of activation functions. We also show the results for J48 as a reference. From these tables, we can see that the results obtained by each type of activation function is different. When using the SegRELU function, we can obtain the good results. It can be explained by the fact that, the characteristic of SegRELU is more adapt than the others when applied for these data sets.

Table 2. The distribution of data sets for training and testing

Data	Normal	Arp	Cafe-latte	Fragmen-tation
O-AWID-Train	5000	500	500	50
O-AWID1-Test	1500	200	200	150
O-AWID2-Test	1000	1000	1000	167
O-AWID3-Test	800	800	800	300
O-AWID4-Test	700	500	500	167

Table 3. Comparison of experimental results of J48 and Deep learning on data set O-AWID1-Test

Method	Acc. (%)	F1-Score (%)
J48	83.04	85.92
CNN-Relu	91.52	91.58
CNN-LRelu	86.86	87.71
CNN-ELU	94.43	94.3
CNN-SegRELU	95.15	96.25

Table 4. Comparison of experimental results of J48 and

Deep learning on data set O-AWID2-Test

Method	Acc. (%)	F1-Score (%)
J48	65.14	62.13
CNN-Relu	85.22	87.37
CNN-LRelu	87.72	89.56
CNN-ELU	55.98	48.76
CNN-SegRELU	88.42	89.43

Table 5. Comparison of experimental results of J48 and Deep learning on data set O-AWID3-Test

Method	Acc. (%)	F1-Score (%)
J48	64.78	61.22
CNN-Relu	89.56	90.55
CNN-LRelu	91.04	93.41
CNN-ELU	61.28	52.45
CNN-SegRELU	92.64	92.07

Table 6. Comparison of experimental results of J48 and Deep learning on data set O-AWID4-Test

Method	Acc. (%)	F1-Score (%)
J48	67.65	63.23
CNN-Relu	96.52	96.15
CNN-LRelu	99.04	98.06
CNN-ELU	72.31	63.61
CNN-SegRELU	98.15	98.08

To summary, the activation function has an important impact for deep learning models. We need more research for developing new activation functions, testing for more kinds of data i.e., image data, intrusion detection data, speech recognition data, etc.

IV. CONCLUSION

This paper makes a brief survey about the main activation functions used for deep learning up to date. The activation function plays an important role in model based on neural network, especially, deep learning. We also apply for the AWID data set. The results obtained show that each activation function contribute different

impact for CNN model. In the near future, the new activation function needs to be further studied. When we deal with real applications, the use of different kinds of activation functions should be tried and choosing the best one.

ACKNOWLEDGEMENTS

The reported study was funded by VAST (VIETNAM ACADEMY OF SCIENCE AND TECHNOLOGY) with project number QTRU01.14/21-22 in 2021.

REFERENCES

1. Yang, E., Joshi, G. P., & Seo, C. (2021). Improving the Detection Rate of Rarely Appearing Intrusions in Network-Based Intrusion Detection Systems. *Computers, Materials Continua*, 66(2), 1647–1663.
2. Bouasker, S., Hamrouni, T., Ben Yahia, S. (2012). New Exact Concise Representation of Rare Correlated Patterns: Application to Intrusion Detection. In: Tan, PN., Chawla, S., Ho, C.K., Bailey, J. (eds) *Advances in Knowledge Discovery and Data Mining. PAKDD 2012. Lecture Notes in Computer Science*, vol 7302. Springer, Berlin, Heidelberg.
3. Khor, KC., Ting, CY. & Phon-Amnuaisuk, S. A cascaded classifier approach for improving detection rates on rare attack categories in network intrusion detection. *Appl. Intell.* 36, 320–329 (2012).
4. C. Kolias, G. Kambourakis, A. Stavrou and S. Gritzalis, "Intrusion Detection in 802.11 Networks: Empirical Evaluation of Threats and a Public Dataset," in *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 184-208, Firstquarter 2016.
5. Vinod Nair and Geoffrey E. Hinton. Rectified linear units improve restricted boltzmann machines vinod nair. *Proceedings of ICML*, 27:807-814, 06 2010.
6. Andrew L. Maas, Awni Y. Hannun, and Andrew Y. Ng. Rectifier nonlinearities improve neural network acoustic models. In *in ICML Workshop on Deep Learning for Audio, Speech and Language Processing*, 2013.
7. D.-A. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs)," *Under Review of ICLR2016* (1997), 11/23 2015.
8. Hong-Seng Gan Viet-Vu Vu, Doan-Vinh Tran, Viet-Thang Vu, Akinobu Shimizu, Muhammad Hanif Ramlee, and Riries Rulaningtyas, SegRELU: A Parameter-Free Activation Function for Deep Neural Network used in Medical Image Segmentation, Submitted.
9. Vu V. T., Pashchenko F. F., Multistage System-Based Machine Learning Techniques for Intrusion Detection in WiFi Network. *Journal of Computer Networks and Communications*, Special Issue *Advances in Machine Learning for Cybersecurity*. – Volume 2019, Article ID 4708201, 13 pages.
10. I. Aljamal, A. Tekeoğlu, K. Bekiroglu and S. Sengupta, "Hybrid Intrusion Detection System Using Machine Learning Techniques in Cloud Computing Environments," 2019 IEEE 17th International Conference on Software Engineering Research, Management and Applications (SERA), 2019, pp. 84-89.
11. Çavuşoğlu, Ü. A new hybrid approach for intrusion detection using machine learning methods. *Appl Intell* 49, 2735–2761 (2019).
12. Megantara, A.A., Ahmad, T. A hybrid machine learning method for increasing the performance of network intrusion detection systems. *J Big Data* 8, 142 (2021).
13. Ansam Khraisat et al., Hybrid Intrusion Detection System Based on the Stacking Ensemble of C5 Decision Tree Classifier and One Class Support Vector Machine, *Electronics*, 9(1): 173, 2020.
14. Chao L, Wen S, Fong C (2015), CANN: An intrusion detection system based on combining cluster centers and nearest neighbors, *Knowledge Based System*, 78: 13-21
15. Elhag S, Fernandez A, Bawakid A, Alshomrari A, Herrera F (2015), On the combination of genetic fuzzy systems and pairwise learning for improving detection rates on Intrusion Detection Systems, *Expert Syst Appl*, 42 (1): 193-202.
16. Can O, Sahingoz OK (2015), A survey of intrusion detection systems in wireless sensor networks, 2015 6th Int. Conf. on Modeling, Simulation, and Applied Optimization (ICMSAO), pp 1-6, IEEE.
17. Buczak AL, Guven E (2016), A survey of data mining and machine learning methods for cyber security intrusion detection, *IEEE Commun Surveys Tutorial*, 18(2): 1153-1176.
18. Meshram A, Haas C (2017), Anomaly detection in industrial networks using machine learning: A roadmap (in Beyerer J, Niggemann O, Kuhnert C, *Machine learning for Cyber Physical Systems: Selected papers from the Int. Conf. ML4CPS 2016*). Springer Berlin Heidelberg, Berlin, pp 65-72.
19. Hope Eke, Andrei Petrovski, and Hatem Ahriz, Handling Minority Class Problem in Threats Detection Based on Heterogeneous Ensemble

Learning Approach. International Journal of Systems and Software Security and Protection 11, 2 (Jul 2020), 13–37.

20. N. Shone, T. N. Ngoc, V. D. Phai and Q. Shi, "A Deep Learning Approach to Network Intrusion Detection," in IEEE Transactions on Emerging Topics in Computational Intelligence, vol. 2, no. 1, pp. 41-50, Feb. 2018.
21. J. Liu and S. S. Chung, "Automatic Feature Extraction and Selection for Machine Learning Based Intrusion Detection," 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation, 2019, pp. 1400-1405.
22. https://github.com/ThangVu-Viet/Russia_Intrusion-Detection-System



Dr. Viet-Thang Vu is lecturer and the head of Information System division at Hanoi University of Industry, Vietnam. His research interests include cyber-security, image processing, computer vision, machine learning, data mining.



Dr. T.T. Quyen Bui is a researcher and a head of Department of Automation Technology, Institute of Information Technology (IOIT), Vietnam Academy of Science and Technology (VAST) since May 2015. Dr. Bui's current research interests include embedded systems, robotics, measurement systems, computer vision, brain-computer interface, IoT.



Assoc. Prof. Dr. Viet-Vu Vu received the B.S. degree in Computer Science from Ha Noi University of Education in 2000, a M.S. degree in Computer Science from Ha Noi University of Technology in 2004, and a Doctor Degree in Computer Science from Paris 6 University in 2011. He is a researcher at Information Technology Institute, Vietnam National University, Hanoi. His research interests include clustering, active learning, semi-supervised clustering, and E-government applications.



Assoc. Prof. Dr. Hong-Seng Gan received his BEng and PhD in Biomedical Engineering from Universiti Teknologi Malaysia in 2012 and 2016, respectively. He is a faculty member of School of AI and Advanced Computing, Xi'an Jiaotong – Liverpool University, Suzhou, China. His research areas focus on machine learning, computer vision and medical image processing.



Manh-Quang Do completed a Master's degree in Information Systems from Hanoi University of Industry in 2022. Currently, he is working as an AI engineer at Giao Hang Tiet Kiem joint stock company. His main research areas are computer vision and natural language processing.



Thanh-Duc Vu received a BS degree of Information Technology from Moskva Technological University in 2008. Since 2009, he's been with the Information Technology Center at Ha Noi University of Industry, Vietnam, where he is now a lecturer. His main research areas are machine learning, data analytic and data science.



Dinh-Lam Pham is a research professor in the Contents Convergence Software Research Institute at Kyonggi University, South Korea. His research interests

include video retrieving system, deep learning, workflow systems, BPM, workflow-supported social and affiliation networks discovery and analysis, process mining, deep-learning-based process predicting.