ISSN 1533-9211

DOI 10.17605/OSF.IO/P7AD6

MACHINE LEARNING BASED INTRUSION DETECTION SYSTEM FOR NETWORK SERCURITY USING SELF-ORGANIZING MAP

SOUNDARA RAJAN D S[1], Dr. T.V. ANANTHAN[2] and Dr. V.N. RAJAVARMAN[3]

[1]Research Scholar, Computer Science and Engineering, Dr. M.G.R Educational and Research Institute, Chennai, India; Email: srisairajha@gmail.com

[2]Professor, Computer Science and Engineering, Dr. M.G.R Educational and Research Institute, Chennai, India; Email: tvananthan@drmgrdu.ac.in

[3]Professor, Dr. M.G.R Educational and Research Institute, Chennai, India.
Email: nrajavarman2003@gmail.com

Abstract

Machine learning techniques are widely used for detecting network attacks at the network level and the host level in a timely and automatic, to develop an intrusion detection system (IDS) for grading. Malicious attacks is always the need for change and scaling solutions, because it occurred during a very large volume, however, a number of challenges will occur. There is a data set that can be publicly available for further research on another malware of cyber security within the community. Network, plays an important role in modern life, has become the network security is an important field of research. Is an important network security technology Intrusion Detection System (IDS) is to monitor the software running on the network and hardware status. Despite the decades of development, the existing IDS, still, to reduce the improved error rate detection accuracy, are faced with the challenge of detecting unknown attacks. To overcome the issues proposed the method is Self-Organizing Map (SOM) Performance of intrusion detection depends mainly on the accuracy. Accuracy intrusion detection is to reduce the error rate; it must be strengthened in order to increase the detection rate. In order to improve performance, different technologies, have been used in recent works. It is the main work of the intrusion detection system for analysing a huge network traffic data. Well-organized classification method, you need to solve this problem.

Keywords: Machine learning, Intrusion Detection System, Self-Organizing Map (SOM), Accuracy, network traffic.

Subject classification codes: include these here if the journal requires them

1. INTRODUCTION

To detect vulnerabilities various security within the Network Intrusion Detection System (NIDSs) organization's network is a tool necessary for the system administrator of the network. The monitor and intrusion of NIDS enters when it is observed from the network equipment of the organization, network traffic is finished, and the analysis has triggered the alert. Based on the NIDS on the basis, I) signatures (misuse) and ii) anomaly detection: According to the intrusion detection method, NIDSs are divided into two categories. In NIDS like this as Snort, a node specification of the attack it has been pre-installed on the NIDS. The actual network traffic marker packets availability from development NIDS. Generation, or over time requiring huge effort, tagging data set from the trace of raw network traffic collected in real time.

- In addition, in order to maintain a variety of user privacy and confidentiality of the network structure of the internal organization, the network administrator does not want to report that can occur in the network intrusion. Various machine learning techniques ADNIDSs, such as artificial neural network (ANN), support vector machine (SVM), Naive Bayes (NB), has been used to develop a random forest (RF) and self-organizing map (SOM) I will.).

2. RELATED WORK

Deep Learning technique is fruitful, have been applied the reason the extraordinary highlights portrayal from non-named a lot of information, at that point, it is to apply the capacity learned of these to the restricted measure of information and directed characterization [1]. It may name and unlabeled information coming from various dispersions. In any case, they additionally should speak with one another. Traffic light frameworks are principally made out of the Internet center and wired/remote heterogeneous spine network [2]. As of late, these bundle exchanged frameworks have encountered unstable development of organization traffic to the fast improvement of the reason correspondence innovation. Existing organization procedure is mind boggling, and not be adequate to adapt to the continually changing organization conditions achieved by an increment in the immense rush hour gridlock [3]. Generally long arrangement deferral and control and the board of complex organizations, on-request virtual organization work administration chain, is brought about by the transitional server farm flexible optical organization supply [4[. To first to tackle the above issues, has planned the stockpile arrangement of pre-sending of assets. Since in the high-dimensional tangible information to the center to tackle a large number of the PC vision-related errands, it is feasible to extricate a significant degree of portrayal, to assemble a smart framework [5].

To multi-phantom neural organization gained from a majority of columns of profound neural organization work, the second level separation complex from the latter is in a tight articulation [6]. To get familiar with the more profound importance has been utilized to work broadly perceived. Yet in addition, a portion of the exploration of how to utilize the data of primary complex to an assortment of activity film to improve the exactness and productivity of acknowledgment have been made. Exhaustive examination of remote sensor organization (WSN) is, IDS the framework utilizing the machine and Deep Learning (DL) arrangements [7].

The basic framework remote sensor networks Deep learning, customary digital actual frameworks, keen, making maintainable, has become a well-known man-made consciousness innovation [8]. As of late, profound learning and it has been broadly utilized in network space. With the assistance of the amazing profound neural organizations, correspondence organizations, to stay away from a disappointment or clog parcel prospects, you can make an astute exchange activity [9]. Nonetheless, in the lone high computational expense and specialized limits and static organization situation, network stream control calculation of the current profound learning base, the following advancement of the necessities of the age of enormous scope dynamic organization it won't meet [10]. Mist radio access organization has been viewed as an advancement innovation that help of IoT administrations utilizing the edge reserving and edge [11].

Computing the off-stacking and asset portion, the current commitment is wasteful, likewise, they, similar to the solitary static correspondence mode, expands the low dormancy administration and F-RAN kindly consider the raised stance high throughput huge difficulties to the interest [12]. To figure out how to Shenzhen Development is hoping to use with a layered instructor learning joined with an administered learning for fine change. This reasonable methodology, profound organization will actually want to see as a very remarkable particular framework than is truly appropriate for learning.

In the interconnection of proceed with an immense sum and worldwide development of the Internet foundation of the information that is produced each day, interruption identification framework dependent on AI is essential to ensure our monetary and public safety It is a section [13]. Shallow learning and profound learning system of the past page, utilize the technique for a solitary learning model for interruption location. Single learning model methodology, the issue to get a handle on the inexorably intricate information dispersion of the attack model may happen. Adaptable information rate and regulator territory network is a correspondence convention broadly utilized for vehicle recognition and control. Notwithstanding, because of the absence of explicit security instruments, unapproved gadgets can get to the installing prying gadget to the vehicle organization. Fundamentally wellbeing noxious interruption into improve the dangers related with security and protection, you can distribute the harmed vehicle. The motivation behind recognizing network interruption, from the utilization of the typical of the Internet, is to recognize the assaults on the Internet [12]. This is a significant piece of the data security framework. Due to the fast advancement of because of an assortment of organization conduct assault design, there is a need to build up a rapid AI interruption identification calculation with a high discovery rate and low bogus alert rate. Organization interruption discovery technique dependent on AI, it has acquired expanding fame. Since the number and intricacy of new assaults has expanded, you should have a viable and brilliant arrangements [10].

As of late, the measure of huge scope studies, and its pervasive presence, to at least security of spillage, in view of the remote LAN signal, has been done to build up a gadget less interruption identification framework [11]. Nonetheless, the current WiFi-based interruption identification framework is, for the most part, truth be told, to restrict its utilization, bogus alerts brought about by pets, languishing. You will actually want to identify the zero-day assault. So they, solo AI innovation is especially alluring a direct result of the interruption location framework, knew, not just obscure assaults. Indoor Intrusion Detection (IID) is a vital innovation for an assortment of key applications [12].

Interruption location is identified with the layers of network protection to forestall hacking and unlawful demonstrations from happening to the resources of the organization. [11]. Notwithstanding, the current interruption location are confronted with a portion of the issues and difficulties that actually influence the recognition execution. Particularly with the quick improvement of the current organization, the measure of organization information and the scale

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/P7AD6

is expanding step by step, the organization is brimming with unlabeled information, it squeezes the information preparing technique for the IDS [12]. Interruption recognition, interruption endeavors, interruption dynamic organization security safeguard estimates that happened during the attack or distinguishing proof interaction in progress. Right now, low location pace of interruption identification strategy, high caution rate and bogus alert rate, constant execution [10]. This is, to accomplish better location execution, requires an enormous number of or complete information.

3. MATERIALS AND METHOD

Intrusion Detection Systems (IDS) are systems for monitoring and analysis data to detect any intrusion into a system or network. The large volume, variety and speed of data generated within a network is a data analysis process, done by traditional techniques to detect very difficult attacks.

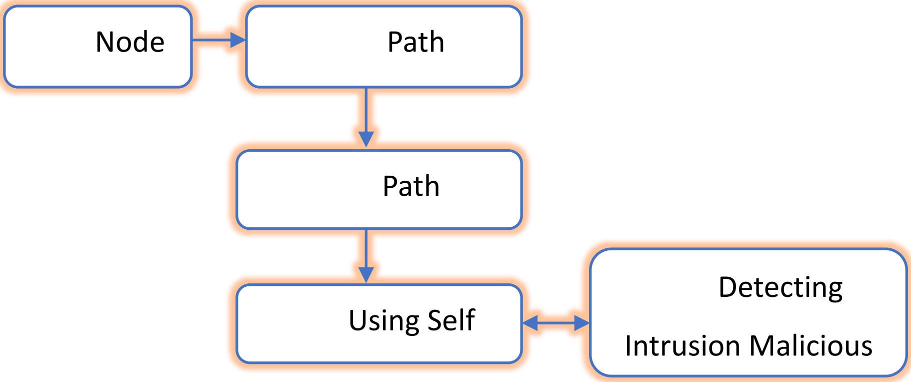

Figure 1: Proposed Block Diagram

Figure 1 describes this model consists of a session-based data preprocessing module and a deep learning architecture module. Network data can successfully reflect the relevance of network data packets.

3.1 Analyzing the Path Verification

When X is X hypothesis label each N to ... N from N, represents a direct path of the sorting node relationships, N∼NJ before Sort: initial network in accordance with the first network relationship, all the nodes is assumed. Since the trust network has a direction, please set in advance the relationship between the expansion nodes.

Step 1: Node Initialization (N=0 ...NJ)

Step 2: Initialize the position for Nodes

$S c = N1 * (Cr + 1) + 1$----- (1)

Step 3: Randomly selected the path to initialize the each node

Step 4: Select the Start and end node

```
    Step 5: Initialize the node position Ni=1 ... N
    Step 6: Each node specified to the node.
    Step 7: Free or ready for route status
    Step 8: For each position Ni
Do
        (I) select the path of Ni (the number of solution)
        (II) Find the next of the route
        (III) Calculate the distance between each node
    Step 9: Calculate node energy values Ni for the solution Xi
    Step 10: End while
    Step 11: End for
```

Where, Ni-position of node, S-source, C-cluster, r- Range, In order to improve the performance, it will be displayed in as late as possible expansion.

3.2 Node Specification

Intrusion detection system (IDS) is an integral part of risk management and overall security architecture, it has become an important security tool. AN IDS is node specification extraction is an important pre-processing step can be thought of as a pattern recognition system. Node specification extraction process, including the selection of the node specification and function. The quality of the selection algorithm of the node specification and function, is one of the most important factors that influence the effectiveness of IDS. Without affecting negatively, the classification accuracy, associated transport function to achieve a reduction in the number of objects greatly improve the overall benefits of IDS. In the practice of intrusion detection, most of the work of building and node specification selection of node specifications are still done manually using the domain knowledge.

3.3 Detection Malicious Attack using Self Organizing Map (SOM)

Intrusion detection system is to adapt to the threat environment that changes from moment to moment, is you should be able to determine the new attack and low false positive, and is a basic network infrastructure of the defense. The researchers, so as to be able to reliably detect the abnormality, has developed various methods of supervision and disciplinary supervision from data mining and machine learning.

Algorithm steps for SOM

```
    Step 1: Initialize packets
    Step 2: WHILE (it does not meet the stop condition)
            FOR p =1 to number of particles
            Select attributes
            Separate Training data and Test data using k-fold cross validation
            Train on Training data
    Step 4: SOM→ Classify using Test data
            Store Intrusion detection rate in an array
            NEXT p
            Update Node position
             NEXT generation until stopping criterion
    Step 5: End.
```

Where, SOM-Self Organizing Map, P-node position, In order to detect the normal record, enter the 1, set it, all the other attack is set to 2. This class, then, are selected with the help of node specifications and binary, classification is done by SOM algorithm. IDS is believed to be the adaptation of the classification, and the particles having the highest IDS in the group is considered the best of the particles. In each generation, processing of attribute selection and classification is done.

4. RESULT AND DISCUSSION

Statistical node specification selection technique, in most cases, by comparison with the levels, identifies the attributes of the data set and assign all attributes based on the level of the associated class. In this way, the attributes of the highest rank is fixed, it is compared with the class label along each attribute from other fixed attributes. In this way, you do not have that any combination of attributes must be compared to the class.

4.1 Analysis of the Security performance

Network security is any measure designed to protect the usefulness and integrity of your network and data. This includes both hardware and software technologies. It targets a variety of threats. It will attack or stop them from spreading in the network. Effective network security manages access to your network.

Security Performance = Process time +drop time/no.of.nodes ------------ (2)

Table 1: Analysis of the Security performance

Num.of.Data(packets)	SVM in %	KNN in %	SOM in %
8	30	36	40
18	39	45	52
28	43	51	57
38	50	59	67
48	62	76	88

Table 1 shows the Security performance analysis based on the collection of packets, in the accuracy improving the proposed method.

ISSN 1533-9211

DOI 10.17605/OSF.IO/P7AD6

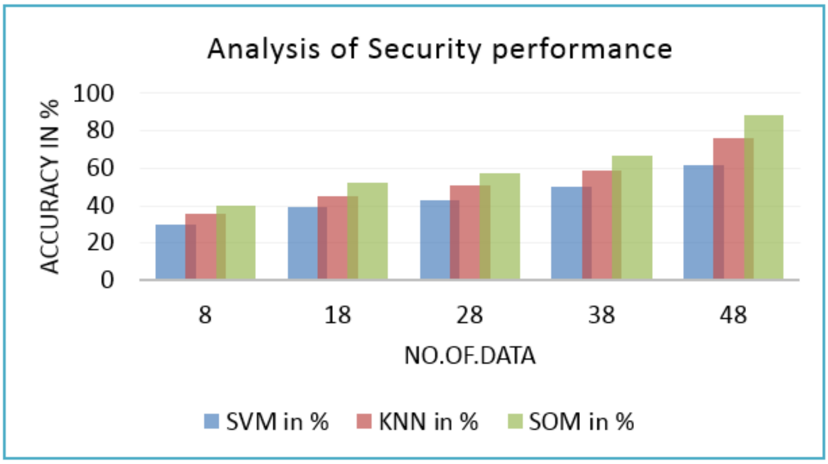

Figure 2: Analysis of Security Performance

Figure 2 shows the Security Performance analysis based on the packets comparing the existing and proposed system, in the proposed method Self Organizing Map (SOM) is 88% accuracy better than previous methods.

4.2 Analysis the False Detection attack

False detection attack estimation of data-driven programs to detect stealth attacks fake data. Many classifiers are used and the results of the nodes, which is also characterized by the unique ensemble of classifiers learning, networks.

Percentage of false Detection attack =FP/Total attacks *100 ------ (3)

FP-False Positive rate.

Table 2: Analysis of False Detection attack

Num.of.Data	SVM in %	KNN in %	SOM in %
8	33	30	27
18	38	35	30
28	44	42	40
38	47	45	39
48	45	40	35

Table 2 shows the false Detection attack analysis based on the Reducing the error rates using the packets, comparing the previous and proposed system.

DOI 10.17605/OSF.IO/P7AD6

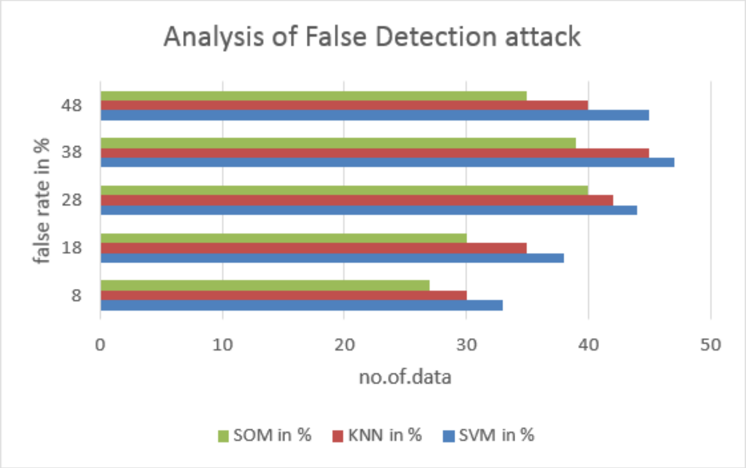

Figure 3: Analysis of False Detection attack

Figure 3 describes the False Detection attack using the packets, in the packets comparing the results in previous and proposed methods, in proposed method reduces the false rate 35% better than previous methods.

4.3 Analysis the Time Complexity

The number of time complex nodes and alternate paths in a network that exists in the computer network transferring the alternate path selection in minimum of time, as well as in a variety of communication nodes and communication protocol.

T (n) = t (n1) + t(n2) + ... + t (n N) -------- (4)

T-Time, N-Node specification

Table 3: Analysis of the Time complexity

Num.of.Data	SVM in sec	KNN in sec	SOM in sec
8	29	25	24
18	34	33	32
28	38	35	34
38	40	37	30
48	44	42	37

Table 3 shows the time complexity reduce based on the packets, in the time is reduced in comparing the previous and proposed methods.

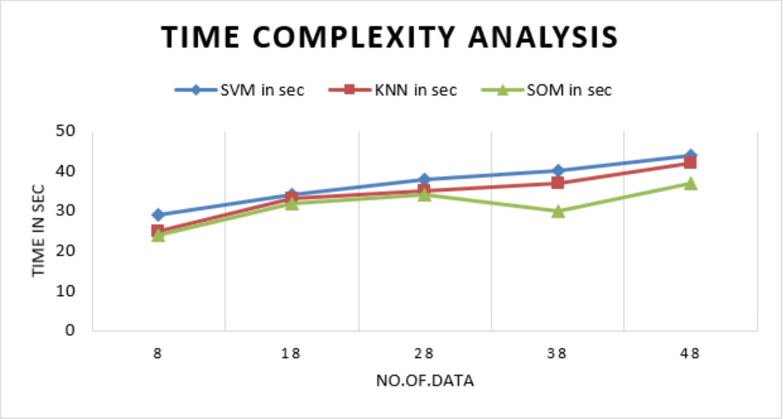

Figure 4: Analysis of time complexity

Figure 4 analysis the time complexity comparing the existing and proposed method, in the proposed method Self Organizing Map (SOM) is 35sec reduced the time.

5. CONCLUSION

Deep neural network, such as intrusion detection, has been proven the most of the machine learning task, in its effectiveness. Unfortunately, a recent study, deep neural networks, image classification, and found that susceptible is an example of a conflict. They are, by introducing subtle changes of the original pixel of the image, leaving the attack to deceive the network to some of the opportunities for misjudgment.

Reference

1. Z. M. Fadlullah et al., "State-of-the-Art Deep Learning: Evolving Machine Intelligence Toward Tomorrow's Intelligent Network Traffic Control Systems," in IEEE Communications Surveys & Tutorials, vol. 19, no. 4, pp. 2432-2455, Fourthquarter 2017, doi: 10.1109/COMST.2017.2707140.
2. B. Li, W. Lu, S. Liu and Z. Zhu, "Deep-learning-assisted network orchestration for on-demand and cost-effective VNF service chaining in inter-DC elastic optical networks," in IEEE/OSA Journal of Optical Communications and Networking, vol. 10, no. 10, pp. D29-D41, Oct. 2018, doi: 10.1364/JOCN.10.000D29.
3. L. Shao, D. Wu and X. Li, "Learning Deep and Wide: A Spectral Method for Learning Deep Networks," in IEEE Transactions on Neural Networks and Learning Systems, vol. 25, no. 12, pp. 2303-2308, Dec. 2014, doi: 10.1109/TNNLS.2014.2308519.
4. X. Chen, J. Weng, W. Lu, J. Xu and J. Weng, "Deep Manifold Learning Combined With Convolutional Neural Networks for Action Recognition," in IEEE Transactions on Neural Networks and Learning Systems, vol. 29, no. 9, pp. 3938-3952, Sept. 2018, doi: 10.1109/TNNLS.2017.2740318.
5. S. Otoum, B. Kantarci and H. T. Mouftah, "On the Feasibility of Deep Learning in Sensor Network Intrusion Detection," in IEEE Networking Letters, vol. 1, no. 2, pp. 68-71, June 2019, doi: 10.1109/LNET.2019.2901792.
6. F. Tang, B. Mao, Z. M. Fadlullah, J. Liu and N. Kato, "ST-DeLTA: A Novel Spatial-Temporal Value Network Aided Deep Learning Based Intelligent Network Traffic Control System," in IEEE Transactions on Sustainable Computing, vol. 5, no. 4, pp. 568-580, 1 Oct.-Dec. 2020, doi: 10.1109/TSUSC.2019.2929935.

7. G. M. S. Rahman, T. Dang and M. Ahmed, "Deep reinforcement learning based computation offloading and resource allocation for low-latency fog radio access networks," in Intelligent and Converged Networks, vol. 1, no. 3, pp. 243-257, Dec. 2020, doi: 10.23919/ICN.2020.0020.

8. S. N. Tran and A. S. d'Avila Garcez, "Deep Logic Networks: Inserting and Extracting Knowledge From Deep Belief Networks," in IEEE Transactions on Neural Networks and Learning Systems, vol. 29, no. 2, pp. 246-258, Feb. 2018, doi: 10.1109/TNNLS.2016.2603784.

9. W. Zhong, N. Yu and C. Ai, "Applying big data based deep learning system to intrusion detection," in Big Data Mining and Analytics, vol. 3, no. 3, pp. 181-195, Sept. 2020, doi: 10.26599/BDMA.2020.9020003.

10. T. Yu and X. Wang, "Topology Verification Enabled Intrusion Detection for In-Vehicle CAN-FD Networks," in IEEE Communications Letters, vol. 24, no. 1, pp. 227-230, Jan. 2020, doi: 10.1109/LCOMM.2019.2953722.

11. K.S.Gautam,Senthil Kumar Thangavel,Video analytics-based intelligent surveillance system for smart buildings, Springer Soft computing ,pp.2813-2837,2019

12. G. Pu, L. Wang, J. Shen and F. Dong, "A hybrid unsupervised clustering-based anomaly detection method," in Tsinghua Science and Technology, vol. 26, no. 2, pp. 146-153, April 2021, doi: 10.26599/TST.2019.9010051.

13. Y. Lin, Y. Gao, B. Li and W. Dong, "Revisiting Indoor Intrusion Detection With WiFi Signals: Do Not Panic Over a Pet!," in IEEE Internet of Things Journal, vol. 7, no. 10, pp. 10437-10449, Oct. 2020, doi: 10.1109/JIOT.2020.2994101.