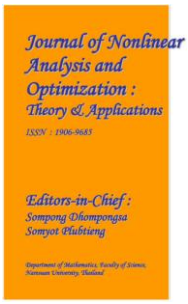
Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 13, Issue. 2 : 2022
ISSN : **1906-9685**



DEEP LEARNING MODEL FOR NETWORK INTRUSION DETECTION SYSTEM UTILIZING CONVOLUTION NEURAL NETWORK

Srinivas Akkepalli, Osmania University, Hyderabad, Telangana, India ,500007.
Sagar K, Sreyas Institute of Engineering and Technology-68, JNTUH, Hyderabad, Telangana, India
sress2020@gmail.com

Abstract Software-defined and organizing Networks(SDN). that isolates the controller from the organize gadgets i.e. switches. The centralized engineering of the SDN encourages the generally organize administration and addresses the necessity of current information centers. Whereas there are more benefits addressed by the SDN architecture, the hazard of un used assaults could be a critical problem and can avoid the wide appropriation of SDNs. The SDN controller may be a pivotal component, and it is an appealing target for the gate crashers. In case the attacker effectively gotten to the SDN controller, it can course the activity based on its possess pre requisites, causing extreme harm to the complete networks. The arrange interruption location frameworks (NIDSs) are vital instruments to distinguish and secure the network environment from pernicious exercises and odd assaults.Deep Learning (DL) has as of late appeared alluring comes about in a assortment of issues, such as content discourse applications, CNN based Regularization handle these issues and effectively detect the instructions.
Keywords:NIDS, Deep learning techniques. NSL-KDD. Handle

1 Introduction

The modern developing SDN organized framework and encourages centralized administration through the programmable control plane, making the Network more adaptable to convey diverse capacities (Jahromi et al., 2018; Elsayed et al., 2021). A few utilize cases within the network world, such as activity designing, arrange observing, quality of benefit (QoS) and datacenter Network have connected in SDN. The adaptable nature of SDN quickens advancement investigate and enhances security measures such as risk discovery and anticipation compared to the customary systems. it has certain security challenges that have to be addressed for wide appropriation of the modern worldview. The imbalance of data leads to miss prediction of intrutions and if attacker succeeds in bringing the controller down, the organize can be uncovered to extreme crashes. The assailant can surge the networkwith the foremost unsafe assaults in SDN such as Refusal of Service (DoS) .
(DDoS)attacks (Elsayed et al., 2020a). Hence, the genuine demands will be denied since the channel transmission capacity and the arrange assets are intensely expended. There are two primary classes of NIDSs based on the discovery approach: signature-based NIDS and anomaly-based interruption location framework (Khraisat et al., 2019), the signature of assaults is put away within the recognize database. In case the observed activity is maped with signature, a caution is generated, referring to the identified assault. the advancement of the anomaly-based IDS may be a essential inquire about activity since the security challenges are Zero attacks detection is still facing aproblem
To handle above issues we Proposed a novel crossover models that's based on a regularizer method,

2.Literature Survey

The researchers try to develop a robust IDS, for this they used ML and Dl models and most of these models detect the malicious attacks.
Li et al. (2018) presented an anomaly detection model to protect the SDN network. This model composed of 2-phases in first phase it used Batalgorithm (BA)to reduce features and extract fine features, next phase contains classification done through Random Forest algorithm.

Gao et al. (2019) presented setting up multiple DT with adaptive voting algorithm. The prediction accuracyimproved and the obtained accuracy for MultiTreeis84.23% andEnsembleVotingis85.2% verified these models with NSL KDD data set.

Jan et al. (2019) presented SVM algorithm to detection malicious attacks in IoT networks. The model is trained using three features derived from the packet arrival rate attribute. The three extracted features are obtained by calculating the mean, median and maximum values of packet arrival rate attribute. it makes the method not a promising solution.

Santos et al. (2020) examined the performance of four various ML-Algorithms, namely SVM, RF, DT and Multiple Layer Perceptron (MLP), against DDoS attacks under the SDN context. The Scapy tool is employed to generate benign and malicious traffic. SVM process large amount of data leads to overfitting problem and exhibit low classification accuracy. The classification of IDS depends on feature extraction methods.. The features that can be used for one attack category may not be suitable for other categories, since the attack scenarios are continuously changing and evolving (Elsayed et al., 2019). the ML techniques provide high performance when the labeled data has a small.

Xiao et al. (2019) presented a DL model with CNN algorithm verified using KDDCup-'99' dataset. The PCA and autoencoder are used in the first stage. The input vector dimensions are reduced from (1 × 122 vector) to 1 × 121 or 1 × 100 dimension reduction vectors. The network feature vectors are converted to image format with 11 × 11 or 10 × 10 matrices, and then the transformed 2-dimensional matrix is passed to the CNN input layer. The CNN model i.e. CNN-IDS, which is based on Lenet-5 typical model, is proposed to extract and analyze the characteristics of the network traffic. This model produced an accuracy of 94.0%, but they failed to achieve a reasonable performance for U2R and R2L attacks (the detection rates of U2R and R2L are 20.61% and 18.96%, respectively).

Andrew Y.Nget al. presented a Regularization techniques to reduce the dimenalitty of feature set and over fitting of data. Two types of Regularization techniques L1 or Lasso regularizer The L1 regularizer penalizes the weight matrix's absolute values from reaching larger values. for less important features, it decreases the weight value to zero and fixed classifier boundaries, these features are not used. Thus, the L1 regularizer is used to select or reduce features.

i. L1 regularize used formula

$\lambda\sum_{i=1}^{n}$ |wi | ----------(1)

Where " λ "is regularization parameter and " n " is the number of features in data set, and "w_I" is the corresponding weight value of i^{th} feature.

ii. L2 regularize or Ridge regularize used to perform less important features weight value does not need to be zero. The characteristics of corresponding coefficient values are reduced and kept greater than zero. The square magnitude values are taken from the weight matrix for this purpose, known as the L2 regularizer .

$\lambda \sum_{i=1}^{n} W\text{i}^2$ ----------(2)

Wherei=1 to n is used for each feature, and " w" is the coefficient value of each feature. "λ" parameter is used to put an extra penalty on the i^{th} weight values. As it controls the magnitude of the coefficient values and here "λ" value selection is crucial task.

Draw back of L1 and L2 Regularizers

L1 regularizer is used for the feature selection or reduction, while L2 gives unimportant features less weight. The major drawback of these regularizers is that they control only individual weight values and do not consider the relationship between entries in the weight matrix.

3.Proposed model

StanderedDiviationReg regularizer We implemented a new regularizer (SD-Reg) to overcome this constraint in L1 and L2 and lthat considers the weight values' dispersion, known as the standard deviation. The SD-Reg regularizer takes the standard deviation of the weight matrix and multiplies it by λ. The motive behind this is to create a weight decay adaptive form. Consequently, the regularizer prevents the learning model from taking widespread values from the weight space and the penalty is equal to 1 for all regularizers ($L1$, $L2$,$elastic\text{-}net$ and the $SD\text{-}Reg$). Further, to maintain the dimensions,the penalty term λ controls the spread of the regularizer by enforcing constraints. Thus, the learning model does not allow to adapt widespread values from weight space.

The mathematical formulation of the new regularizer is

$\lambda\sigma(w)$ ---------- (3)

Where "σ" denotes the standard deviation of weight values givem as

$$\sigma(w) = \sqrt{\frac{1}{nk}\{\sum_{i=1}^{nk} w_i^2} \quad - \frac{1}{nk}\left(\sum_{i=1}^{nk} w_i\right)2 \text{ ----------}\{4)$$

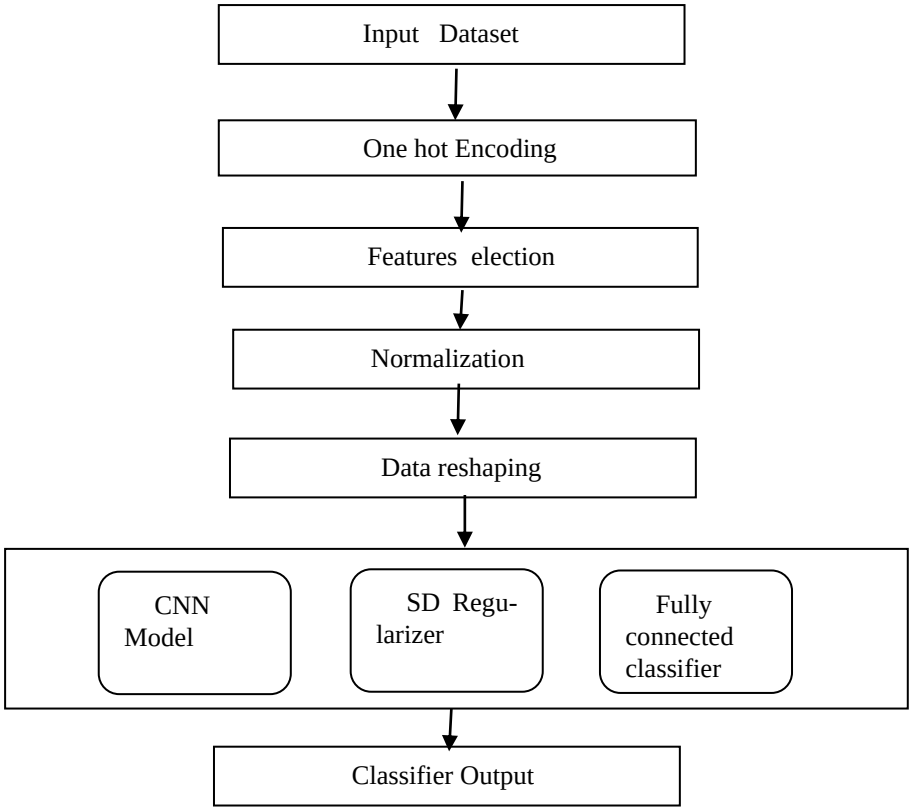
k is the number of rows in the weight matrix and i is the ith row of the weight matrix. The parameter "λ"is used to control the values of the weight matrix, and n is the number of columns in each i^{th} row of the weight matrix. So, n is the size of the weight vector.

Theminimized loss function in our methodology formula:

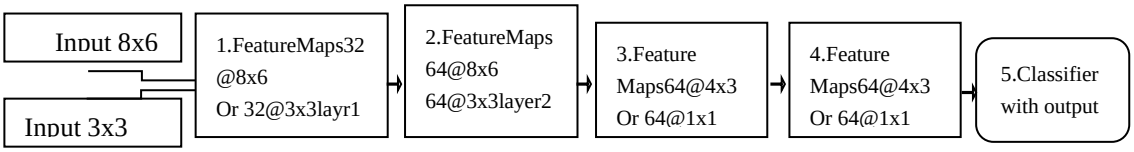
$min_w\{f(X, y{:}w) + \lambda\sigma(w)\}$----------(5)

Thus, we minimize the loss function concerning "w" by a standard deviation to adopt values within a specific range.

Proposed Model FlowChart:



Proposed Model:



[3x3convolution_1+ReLU,Regularization],[3x3convolution_2+ReLU],[2x2 pooling],[Dropout]

{----------Feature-Selection----------}

The proposed model involves CNN architecture combine with the ML algorithms (SVM, KNN, and RF). The CNN is employed to extract the deeper representations of the data features and the classification done by using ML algorithms.

Proposed model implementation: At the first stage, fit the input data for the DL model. CNN takes input data in the form image data. By converting the network data in the form of non-image data into an image, CNN convert the input data from one-dimensional into a two-dimensional matrix. The dimension of the input image is 8 × 6 or 3 × 3 for a subset of 48 and 9, respectively.

The considered hyper-parameters for our CNN model. The number of convolution layers depends on the properties of the input image. we implement two convolution layers. First layer output dimen-

sion is 32 and Second convolution layer output dimention is 64.The filter of 3 × 3 size is used for each layer with a stride equal to 1. Another max-pooling layer of 2 × 2 size and stride of 1 followed the second convolution layer. While each convolution layer learns the feature representation of the previous output, the pooling layer minimizes the dimensions of the feature map. The output from the pooling layer is reshaped (flatten layer) for a fully connected layer with a number of neurons equal to 128. The nonlinear mapping function Relue is used for all layers before the output layer. The classification layer is used to classify the input traffic into normal or attack class.

We implement a set of experiments in the first experiment, the SoftMax activation function with various regularization methods is used to classify the output features produced from the lower convolution layers. We compare the performance of the SD-Reg with L1 and L2 regularize methods. In the second experiment, we test ML Modele by replacing the SoftMax layer and those modele is SVM,. The experimental results proves that the SD-Reg regularize method provides a high performance than the old L1 and L2 methods.

Hyper-parameters:

Hyper-parameters	Optimal values
Convolutional layers	2
Number of filters	32, 64
Kernel size	3 × 3
Stride	1
Pooling layer	Max (2 × 2)

3.1. Experimental setup

The experiment was designed and executed using Python programming language, where Keras with Tensorflow backend library is used for all proposed approaches.

3.1.1.Experimental Environment

Operating System Windows 10 pro 64-bit , Memory 64 GB

CPU Intel(R) UHD Graphics 620,17-865Ou @1.90 GH(8 Cores)

Anaconda 4.9.2 python 3.7.0 keras 2.4.2

Tensorflow 2.2.0 2.2.0

3.2. The evaluation metrics

We used the most common performance measures like the accuracy, precision, recall, and F-score metrics to evaluate the performance of all models, as the following equations.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{----------(4)}$$

$$Precision = \frac{TP}{TP + FP} \quad \text{----------(5)}$$

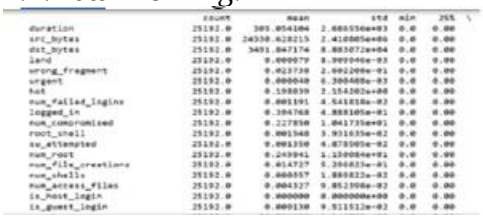
$$Recall = \frac{TP}{TP + FN} \quad \text{----------(6)}$$

F-score = 2 × (Precision × Recall)/(Precision + Recall) --------- (7)

4. Experimental results and analysis:

This section provides a detailed analysis of the results obtained using our proposed models. The performance of the proposed models is evaluated on the NSL-KDD dataset by conducting a experiment describes the experimental configuration used to evaluate the model parameters.

4.1.Data Profiling:



Overall comparison of Results

Work	Detection Model	FS	Precision(%)		Recall(%)		F1 Score(%)	
			Normal	Anomaly	Normal	Anomaly	Normal	Anomaly
1.Alanoud Alsaleh et al.(2021)	SSA XG-boost	SSA		98.0		99.1		99.0
2.Andrew Y. Ng et al.	CNN-Softemax (L1	CNN	99.04	97.38	94.69	99.81	97.08	98.58
3. .Andrew Y. Ng et al.	CNN-Softemax (L2)	CNN	99.43	97.65	95.25	99.72	97.30	98.68
4.Our proposed model	CNN-Softemax (SD-Reg)	CNN	99.82	98.67	96.515	98.28	98.51	98.32
5.proposed model	CNN-SVM	CNN	97.89	95.75	91.40	98.99	94.72	97.82

1. Conclusion

The SD-Reg method used to overcome the overfitting issue of the classifier models.Our proposed model SD-Reguleraization outperformed and compared results with exising modeles L1 and L2 in tems precission ,Recall,F1-score.CNN with SD regularization and soft Max anomaly precission=98.67,Anomaly Rcall=98.28,F1Score=98.32.

In future Deep learning models with optimal feature selection lead to improve anomaly detection rate.

References

1. The Influence of Salp Swarm Algorithm-Based Feature Selection on Network Anomaly Intrusion Detection, Alanoud Alsaleh AndWojdanBinsaeedan, Digital Object Identifier 10.1109/ACCESS.2021.3102095

2. Said Elsayed, M., Le-Khac, N.-A., Dev, S., Jurcut, A.D., 2020. Network anomaly detection using LSTM based autoencoder. In: Proceedings of the 16th ACM Symposium on QoS and Security for Wireless and Mobile Networks. pp. 37–45..

3. Ng, A.Y., 2004. Feature selection, L1 vs. L2 regularization, and rotational invariance. In: Twenty-First International Conference on Machine Learning - ICML '04. ACM Press, http://dx.doi.org/10.1145/1015330.1015435.
4. Khraisat,A., Gondal, I., Vamplew, P., Kamruzzaman,J.,2019. Survey of intrusiondetection-systems:Techniques,datasetsandchallenges.Cybersecurity2(1),1–22.
5. Kim,J.,Kim,J.,Kim,H.,Shim,M.,Choi,E.,2020.CNN-
Santos,R.,Souza,D.,Santo,W.,Ribeiro,A.,Moreno,E.,2020.Machinelearningalgorithms to detect DDoS attacks in SDN. Concurr. Comput.: Pract. Exper. 32(16), e5402.
6. oupas, P., Chamou, D., Giannoutakis, K.M., Drosou, A. and Tzovaras, D., 2019, December. An intrusion detection system for multi-class classification based on deep neural networks. In 2019 18th IEEE International conference on machine learning and applications (ICMLA) (pp. 1253-1258). IEEE.
7. Hindy, H., Atkinson, R., Tachtatzis, C., Colin, J.N., Bayne, E. and Bellekens, X., 2020. Utilising deep learning techniques for effective zero-day attack detection. Electronics, 9(10), p.1684.
8. Lin, P., Ye, K. and Xu, C.Z., 2019. Dynamic network anomaly detection system by using deep learning techniques. In Cloud Computing–CLOUD 2019: 12th International Conference, Held as Part of the Services Conference Federation, SCF 2019, San Diego, CA, USA, June 25–30, 2019, Proceedings 12 (pp. 161-176). Springer International Publishing.
9. Selvakumar, B. and Muneeswaran, K., 2019. Firefly algorithm based feature selection for network intrusion detection. Computers & Security, 81, pp.148-155.
10. Prasad, M., Tripathi, S. and Dahal, K., 2020. Unsupervised feature selection and cluster center initialization based arbitrary shaped clusters for intrusion detection. Computers & Security, 99, p.102062.
11. RM, S.P., Maddikunta, P.K.R., Parimala, M., Koppu, S., Gadekallu, T.R., Chowdhary, C.L. and Alazab, M., 2020. An effective feature engineering for DNN using hybrid PCA-GWO for intrusion detection in IoMT architecture. Computer Communications, 160, pp.139-149.
12. Ashiku, L. and Dagli, C., 2021. Network intrusion detection system using deep learning. Procedia Computer Science, 185, pp.239-247.
13. Lin, P., Ye, K. and Xu, C.Z., 2019. Dynamic network anomaly detection system by using deep learning techniques. In Cloud Computing–CLOUD 2019: 12th Interna-tional Conference, Held as Part of the Services Conference Federation, SCF 2019, San Diego, CA, USA, June 25–30, 2019, Proceedings 12 (pp. 161-176). Springer In-ternational Publishing.
14. Vinayakumar, R., Alazab, M., Soman, K.P., Poornachandran, P., Al-Nemrat, A. and Venkatraman, S., 2019. Deep learning approach for intelligent intrusion detec-tion system. Ieee Access, 7, pp.41525-41550.
15. Almiani, M., AbuGhazleh, A., Al-Rahayfeh, A., Atiewi, S. and Razaque, A., 2020. Deep recurrent neural network for IoT intrusion detection system. Simulation Modelling Practice and Theory, 101, p.102031.
16. Laghrissi, F., Douzi, S., Douzi, K. and Hssina, B., 2021. Intrusion detection sys-tems using long short-term memory (LSTM). Journal of Big Data, 8(1), p.65.
17. Wang, M., Lu, Y. and Qin, J., 2020. A dynamic MLP-based DDoS attack detec-tion method using feature selection and feedback. Computers & Security, 88, p.101645.
18. Maniriho, P., Niyigaba, E., Bizimana, Z., Twiringiyimana, V., Mahoro, L.J. and Ahmad, T., 2020, November. Anomaly-based intrusion detection approach for IoT networks using machine learning. In 2020 international conference on computer engineering, network, and intelligent multimedia (CENIM) (pp. 303-308). IEEE.
19. Krishnaveni, S., Sivamohan, S., Sridhar, S.S. and Prabakaran, S., 2021. Efficient feature selection and classification through ensemble method for network intrusion detection on cloud computing. Cluster Computing, 24(3), pp.1761-1779.
20. Karthikeyan K. R. and A. Indra, "Intrusion Detection Tools and Techniques –A Survey," International Journal of Computer Theory and Engineering vol. 2, no. 6, pp. 901-906, 2010.

21. Bhargavi R. and V. Vaidehi, "Complex Event Processing for Object Tracking and Intrusion Detection in Wireless Sensor Networks," International Journal of Com-puter Theory and Engineering vol. 3, no. 3, pp. 435-439, 2011.
22. Maheyzah Md Siraj, Mohd AizainiMaarof and Siti Zaiton Mohd Hashim, "A Hybrid Intelligent Approach for Automated Alert Clustering and Filtering in Intrusion Alert Analysis," International Journal of Computer Theory and Engineering vol. 1, no. 5, pp. 539- 545, 2009.
23. Swimpy Pahuja and Anita Singhrova, " Preventive Alternate Path Routing Algo-rithm against Intrusion in Sensor Area Network," international journal of computer theory and engineering vol. 5, no. 2, pp. 188-191, 2013.
24. M. Mahboubian and Nor I. Udzir, "A Naturally Inspired Statistical Intrusion Detection Model," International Journal of Computer Thery and Engineering vol. 5, no. 3, pp. 578- 581, 2013.
25. Sasanka Potluri and Christian Diedrich, "Deep Feature Extraction for multi-Class Intrusion Detection in Industrial Control Systems," International Journal of Computer Theory and Engineering vol. 9, no.5, pp. 374-379 , 2017.