World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(RESEARCH ARTICLE)

Check for updates

A method for improving packet detection accuracy using artificial neural networks in computer networks

Chon RyongIl *, An SongIl, Pak CholRyong, Kim DongKuk, Choe Kang and Ri Chol Ryong

Faculty of Information Science and Technology, Kim Chaek University, Pyongyang, Democratic People's Republic of Korea.

World Journal of Advanced Research and Reviews, 2026, 30(02), 1695-1701

Publication history: Received on 12 April 2026; revised on 18 May 2026; accepted on 20 May 2026

Article DOI: https://doi.org/10.30574/wjarr.2026.30.2.1427

Abstract

This paper proposes an effective method for detecting malicious network data by combining sparse-response deep belief network (SR-DBN) and support vector machine (SVM). SR-DBN is an efficient unsupervised learning model that extracts redundant-free data feature representations, while SVM is adopted to construct a classifier with strong generalization capability in the feature space through supervised training. In this method, SR-DBN is used to realize feature representation of abnormal network payloads, and SVM is applied to complete the classification of normal and abnormal payloads.

Simulation and experimental results demonstrate that the proposed network anomaly detection system achieves a higher detection rate than the stacked autoencoder-based multilayer perceptron.

Keywords: Network Intrusion Detection System (NIDS); Deep Learning (DL); Sparse-response Deep Belief Network (SR-DBN)

1. Introduction

As more people use the Internet for personal or business reasons, different cyber-attacks and intrusions are ever-growing. IDS is one of the most essential consideration of cyber-security. IDS is utilized to recognize successful violations even after they have happened [1].

In [2], Alex Shenfield et al, proposed an offline approach for detecting shellcode patterns within a various file data using artificial neural networks.

V. Kanimozhi and his co-authors stated that their proposed system using ANN can be applied to conventional network traffic analysis, cyber-physical system traffic analysis and can also be applied to the real-time network traffic data analysis [3].

[8] proposed a multi-layer perceptron pretrained by stacked auto-encoders for efficient detection. In this paper we propose a method to detect anomalous payload by combination of SR-DBN and Support vector machine on the network traffic.

The results show that this classification approach is capable of detecting anomalous payload with extremely high accuracy.

* Corresponding author: Chon RyongIl

The rest of this paper is organized as follows: In section II, a background of intrusion detection systems, SR-DBN and Support vector machine is provided, and in section III, the problem domain and proposed method are described and de the experimental results are illustrated. In section IV, the contribution and future work of the paper.

2. Background and previous work

2.1. Intrusion Detection Systems

The term intrusion detection system was first used by James Anderson [4] in the late 70s and early 80s. He introduced the concept of misuse detection and predefined events and provided the basic for future IDS design and development. An IDS is software or hardware designed to detect any malicious activity or attack against the system or network. An IDS collects data from different sources within a computer or a network such as system command, system log, system accounting, security log and network log. Then, it analyzes them to identify possible security violation, and finally, it issues an alert to the system administrator to deal with the intrusion.

Swathi Pai M, Ashoor et al. [5, 6] summarized IDS Functions as: monitoring and analyzing both the user and system activities, analyzing system configurations and vulnerabilities, evaluating the system and file integrity, recognizing patterns of typical attacks, analyzing of abnormal activity patterns and tracking user policy violations.

- There are two main types of IDS: Network-based IDS and Host-based IDS [7].
- NIDS is placed along a network to monitor all network traffic [7].
- HIDS placed on a host to scan and monitor all host processes or devices on the network [7].
- IDS can be further categorized into signature and abnormally based systems.

Signature-based systems store attack pattern data in signature database to compare the intruded data and judge when they are identical each other. The advantage is that it has a high efficiency of detecting attacks listed in pattern database, while disadvantage is that it's difficult to cope with unknown attacks or well-modified known attacks. Moreover, it needs an expert to keep updating attack database.

Unlike signature-based systems, the above ones, on the basis of taking statistical feature for normal behavior into account, consider any feature to be anomalous once it is different from normal behavior. Its advantage is that it is capable of detecting unknown attacks. However, it is difficult to get statistical feature on normal behavior and it has a high percentage of false positive alert. To be worse, it fails to give an alert on anomalous thing.

2.2. Artificial Neural Networks

Artificial Neural Network(ANN) is a very powerful tool to deal with signal processing, computer vision and other classification and regression problems [8].

ANN inspired by the biological neural network of human brain, is based on a set of algorithms to extract high-level abstract features from input data by multiple processing layers and it can automatically infer rules for expected results [9-14].

Deep Neural Network(DNN) which has so many complex neurons and layers endowed with the function of feature extractor that can express the human's recognition diagnosis efficiently in its model, is relatively a very complicated ANN[13,14].

DNN is large set of algorithms which has the function of extracting features for recognition automatically and its architecture is different according to applying object [15].

The Sparse-response deep belief networks are developed on the basis of rate distortion theory, which encodes the original data with as few bits as possible.

Assuming that $\mathbf{V} \in R^n$ is input data, $\mathbf{Z} \in R^m$ is its representation or code of the $\mathbf{V}$, $\mathbf{h}$ is the hidden variables of belief network, $\boldsymbol{P}(\mathbf{h} = \mathbf{1}|\mathbf{v})$is activation probability of $\mathbf{h}$, the K-L distribution between the data probability distribution and p_θ^∞ defined by RBM is $KL(p^0 \parallel p_\theta^\infty)$, according to the rate distortion theory, the dual form of sparse-response RBM is as follows.

$$\cdot \quad \cdot \min_{\{w_{ij},c_i,b_j\}} KL(p^0 \parallel p_\theta^\infty) \qquad (1)$$

Constraint:

$$\cdot \sum_{l=1}^{m} \parallel p(h^{(l)}|v^{(l)}) \parallel \leq \eta \cdot \qquad (2)$$

The training of SR-DBN is based on greedy layer-by-layer manner.

The Support vector machine constructs a hyper-plane of high generalization ability in the feature space and its formalization is as follows.

$$Q(\boldsymbol{w}, b, \xi) = \frac{1}{2} \parallel w \parallel^2 + C \sum_{l=1}^{M} \xi_i : min \cdot \qquad (3)$$

Constraint:

$$\cdot \quad \cdot \ y_i(w^T \boldsymbol{g}(\boldsymbol{x}_i) + b) \geq 1 - \xi_i \ (\xi_i \geq 0, i = \overline{1, M}) \cdot \qquad \cdot \qquad (4)$$

Where y_i is classification label and $\mathbf{g}(\mathbf{x})$ are nonlinear vector mapping by the kernel function.

3. Anomalous payload detection in complex network traffic

3.1. Problem Domain

In general, attack patterns used attackers (such as Buffer Overflow and SQL Injection) are manifested through data segments in network packets.

Though the patterns of the above-mentioned attacks are well-known, well-modified variants can bypass intrusion detection systems, enabling attackers to avoid easy detection.

For example, due to low code complexity, small payload size and highly modified attack approaches, network-based intrusion detection faces numerous challenges when identifying shellcode within complex network traffic.

This causes signature-based detection to generate false-positive alerts.

Therefore, developing and researching intelligent intrusion detection methods is one of the major research trends.

To this end, this paper proposes a method for identifying anomalous network packets with a low false-alert rate on network by the combination of SR-DBN and Support Vector Machine.

3.2. Proposed method

In this paper, a new methodology, in which the feature representation is performed by SR-DBN and the classification based on features is done by Support Vector machine, is proposed as shown in Figure 1.

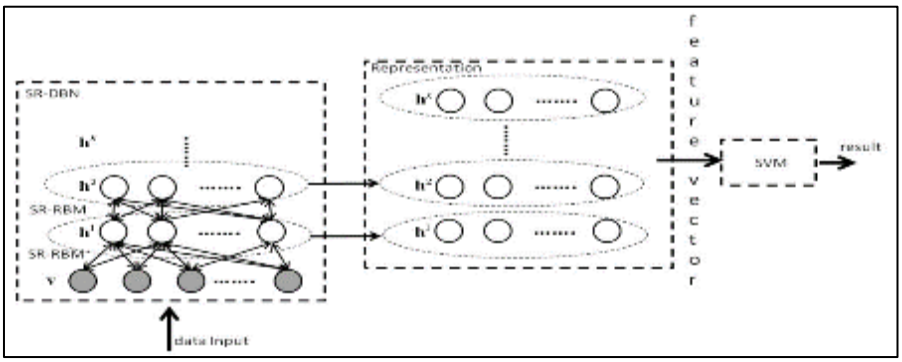

Figure 1 Detection by SR-DBN+SVM

Firstly, the SR-DBN is trained to learn feature representation of normal and abnormal packet data. And then, the feature vector of the packet is constructed based on the hidden outputs of SR-DBN followed by the construction of support vector machine for judging normal or abnormal and the final detection of it.

The detail is as follows:

- Step 1: Collect the data corresponding to normal or attack packets and divide it into training and test dataset.
- Step 2: Define the maximum byte size N of the traffic, for the lack of length of data division is padded by zero padding in order to fix the size of the data.
- Step 3: Construct a SR-DBN with K layers and m neurons in each layer.

Determine the normalization parameter λ and learning rate η for training of the network.

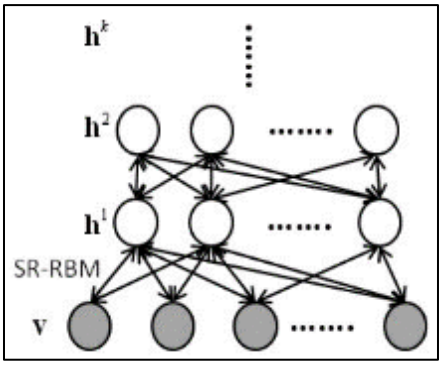

Figure 2 The construction of SR-DBN

Step 4: Train the first SR-RBM shown in Figure 3 by inputting every byte of the traffic into one of the visual units of SR-DBN.

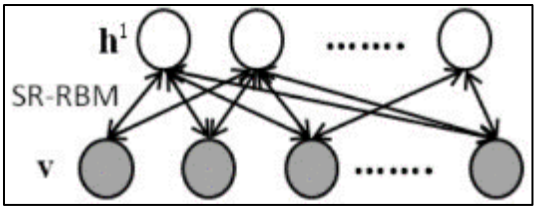

Figure 3 First SR-RBM

Step 4-1: Randomly initiate the weight w_{ij} between the visual and hidden units and thresholds c_i, b_j within visual and hidden units.

Step 4-2: Update the parameters by CD training algorithm.

$$w_{ij} = w_{ij} + \eta(< v_i h_j >_{p^0} - < v_i h_j >_{p^1}) \quad (5)$$

$$c_i = c_i + \eta\left(< v_i >_{p^0} - < v_i >_{p^1}\right) \quad (6)$$

$$b_i = b_i + \eta(< h_i >_{p^0} - < h_i >_{p^1}) \quad (7)$$

, where $<\cdot>_{p^1}$ is calculated from the data reconstructed after a period of Gibbs Sampling.

Step 4-3: Update normalization parameters.

$$w_{ij} = w_{ij} + \lambda\eta \sum_{i=1}^{m} p_j^{(l)}(1 - p_j^{(l)})\, v_i^{(l)} \quad (8)$$

$$b_i = b_i + \lambda\eta \sum_{i=1}^{m} p_j^{(l)}(1 - p_j^{(l)}) \quad (9)$$

,where

$p_j^{(l)} = sigmoid(\sum_{i=1}^{m} v_j^{(l)} w_{ij} + b_j)$ (10)

Step 4-4: Repeat Step 4-2 and 4-3 until convergence.

Step 5: Calculate the hidden layer responses on the training samples while fixing trained parameters of the first SR-RBM and use them as the visual input data of the second SR-RBM shown in Figure 4.

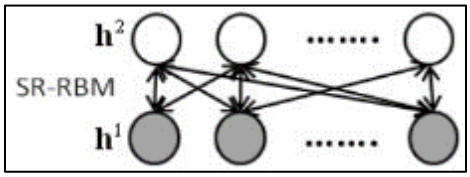

Figure 4 Second SR-RBM

Step 6: Get parameters of second SR-RBM through Step 4.

Step 7: Repeat Step 5 and Step 6 on all K layers.

Step 8: Prepare the training dataset again with the output vectors of the hidden layer on every training samples while fixing all the trained parameters.

Step 9: For every training samples, allocate them labels $y_i = \{1, -1\}$ respectively according to whether it is abnormal or not.

Step 10: Determine the kernel function of SVM and parameter C and calculate the solution of the following optimization problem.

$$L(\alpha) = \sum_{i=1}^{M} \alpha_i - \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} \alpha_i \alpha_j y_i y_j H(\boldsymbol{x}_i, \boldsymbol{x}_j) \quad (11)$$

Constraint:

$$\sum_{i=1}^{M} y_i \alpha_i = 0 \; (0 \le \alpha_i \le C, i = \overline{1, M}) \quad (12)$$

Step 11: Evaluate the network on the test dataset.

The classification function on the feature representation of the unknown packet through the trained SR-DBN is as follows:

$$\cdot \; f(\mathbf{x}) = \sum_{i \in S} \alpha_i y_i \, H(\mathbf{x}_i, \mathbf{x}) + b \quad (13)$$

$$b = \frac{1}{|U|} \sum_{i \in U} \left(y_i - \sum_{j \in S} \alpha_i y_i \, H(\boldsymbol{x}_i, \boldsymbol{x}) \right) \quad (14)$$

, where S -index set of support vectors

U -index set of unbounded support vectors

|U|-the number of elements of set.

For any unknown sample $\mathbf{x}$, if $f(\mathbf{x}) \geq 0$ it is normal and if $f(\mathbf{x}) < 0$, is classified as abnormal.

3.3. Evaluation

For experiment, collected byte data of data division of traffic are input into the input layer of SR-RBM.

Data for training and test of SR-DBN are collected by network traffic for 5 days.

The results are shown in Table 1, where the maximum size of each sample is 1600 bytes.

Table 1 Description of dataset

Day and property	Contents
Monday/1.4G	Normal Traffic
Tuesday/1.1G	Attack Traffic + Normal Traffic
Wednesday/1.3G	Attack Traffic + Normal Traffic
Thursday/1.3G	Attack Traffic + Normal Traffic
Friday/1.15G	Attack Traffic + Normal Traffic

Anomalous data comes from Kali and Acunetix Web Vulnerability Scanner.

Lack of the length of data division is padded by zero padding, which is input into SR-RBM as a type of input data.

In the experiment XEON E3 1210 v6 is used and the size of its hard disk is RAID 5,4 TB.

To generate 10000 training and 5000 test samples, we applied the k-means algorithm on data in table 1 for k = 15000 per class.

Therefore, the dataset contains 20000 training and 10000 test samples. The proposed method is compared with the multi-layer perceptron with stacked auto-encoder as in [16,17]. Two methods are compared in terms of error rate on training and test samples respectively.

The number of units in the input and hidden layer is 1600 and 160 respectively for both stacked auto-encoder and sparse response deep belief network.

For the training of RBM, the learning rate was set to 0.001, and the size of mini-batch was set to 100.

The error rate results of the proposed methods with one hidden layer and various normalization parameters are shown in table 2.

Table 2 The error rate with various λ

λ	Training error rate	Test error rate
0.01	0.015	6.125
0.02	0.038	5.121
0.03	0.024	5.311
0.04	0.036	2.221
0.05	0.042	3.321
0.06	**0.012**	**1.443**
0.07	0.022	5.314
0.08	0.033	1.985
0.09	0.032	3.323

Then, the performance of the proposed method with different numbers of hidden layers and that of the stacked auto-encoder-based method are shown in Table 3, respectively. As presented in Table 3, the proposed method achieves high efficiency; in particular, it attains an extremely high detection rate when the number of hidden layers is 3.

4. Conclusion and future work

This thesis proposes an intelligent intrusion detection system that integrates SR-DBN and SVM to distinguish normal and abnormal network traffic, thereby enhancing the detection performance against anomalous intrusions.

For SR-DBN models with different structural configurations, training and testing were conducted using the datasets listed in Table 1. The corresponding experimental results are presented in Table 3, which verifies the high efficiency of the proposed algorithm.

This study puts forward an offline classification method to identify normal and abnormal network traffic data.

Furthermore, the proposed method can be applied to FPGA-based online network anomaly detection in intrusion detection systems.

In future research, on the basis of deep learning-based abnormal traffic identification, abnormal network behaviors will be further categorized into multiple typical attack types.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed

References

[1] Akbar, S., T.S. Rao, and M.A. Hussain, A Hybrid Scheme based on Big Data Analytics using Intrusion Detection System. Indian Journal of Science and Technology, 2016. 9(33).

[2] Alex Shenfield, David Day, and Aladdin Ayesh, "Intelligent intrusion detection system using artificial neural networks" ICT Express, June 2018.

[3] V. Kanimozhi and T.P. Jacob, "Artificial Intelligence based Network Intrusion Detection with hyper-parameter optimization tuning on the realistic cyber dataset CSE-CIC-IDS2018 using cloud computing", ICT Express, 2019.

[4] Anderson, J.P., Computer security threat monitoring and surveillance. 1980, Technical report, James P. Anderson Company, Fort Washington, Pennsylvania.

[5] Swathi Pai M., B.B.K., Big Data Security Analytic: A classification technique for Intrusion Detection System. ResearchGate, 2015.

[6] Ashoor, A.S. and S. Gore, Importance of intrusion detection system (IDS). International Journal of Scientific and Engineering Research, 2011. 2(1): p. 1-4.

[7] Soniya, S.S. and S.M.C. Vigila. Intrusion detection system: Classification and techniques. in Circuit, Power and Computing Technologies (ICCPCT), 2016 International Conference on. 2016. IEEE.

[8] Shuai Li, Ken Choi and Yunsik Lee, Artificial Neural Network Implementation in FPGA: A Case Study, ISOCC, 297-298, 2016.

[9] L.Deng, D. Yu, "Deep Learning: Methods and Applications", Foundations and Trends in Signal Processing 7:3-4,2014[10] Y.Bengio, "Learning Deep Architectures for AI", Foundations and Trends in Machine Learning 2(1):1-127,2009.

[10] Y. Bengio, A. Courville, P.Vincent, "Representation Learning: A Review and New Perspectives", IEEE Transactions on Pattern Analysis and Machine Intelligence 35(8): 1798-1828,2013.

[11] J. Schmidhuber, "Deep Learning in Neural Networks": An Overview", Neural Networks 61: 85-117,2015

[12] Y.Bengio, Y. LeCun, G. Hinton, "Deep Learning", Nature 521: 436-444,2015

[13] I. Arel, D. C. Rose, T. P. Karnowski, "Deep Machine Learning"- A New Frontier in Artificial Intelligence Research", IEEE Computational Intelligence Magazine, 2013.

[14] I. Goodfellow, Y. Bengio, and A. Courville, "Deep Learning", 2016, MIT Press: Cambridge, MA.

[15] Nan-Nan Ji, Jiang-She Zhang, Chun-Xia Zhang, ELSEVIER, A sparse-response deep belief network based on rate distortion theory, Pattern Recognit. 47(2014) 3179-3191.

[16] Shigeo Abe, "Support Vector Machines for Pattern Classification", Springer Science + Business Media, 15-77,2005.

[17] William Hardy, Lingwei Chen, Shifu Hou, Yanfang Ye, and Xin LiD, L4MD: A Deep Learning Framework for Intelligent Malware Detection, National Science Foundation