Deep Learing Approach for Intelligent Intrusion Detection System

Kaoruko Kawamoto

Keio University

May 22, 2023

Thesis Information

- Title: Deep Learing Approach for Intelligent Intrusion Detection System
- R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat and S. Venkatraman
- Date: Accepted January 3, 2019,
- Journal: IEEE Access, vol. 7, pp. 41525-41550, 2019

Abstract

- IDS system that detect cyberattacks has been developed by using machine learning techniques.
- In this paper, a deep neural network (DNN), is explored to develop a flexible and effective IDS to detect and classify unforeseen and unpredictable cyberattacks.
- Through a rigorous experimental testing, it is confirmed that DNNs perform well in comparison with the classical machine learning classifiers.

Contents

1 Introduction

IDS

An IDS is a proactive intrusion detection tool used to detect and classify malicious access automatically in a timely manner.

- NIDS : network-based intrusion detection system such as firewalls, rooters
- HIDS : host-based intrusion detection system such as anti-virus software

- By combining both NIDS and HIDS collaboratively,an effective deep learning approach is proposed
- In order to capture the contextual and semantic similarity and to preserve the sequence information of system calls, NLP are explored

2 Stages of Compromise

Phases of Cyberattacks

1. reconnaissance : an attacker tries to collect information related to hosts and services
2. exploitation : an attacker utilizes a particular service with the aim to access the target computer
3. reinforcement : an attacker follows camouflage activity and then installs supplementary tools and services to take advantage of the privileges gained
4. consolidation : An attacker obtains a complete control over the system
5. pillage : an attacker steals confidential data and CPU time, and launches an impersonation attack

3 Related Works

NIDS

- weakness : high FPR and high false negative alerts
- solution : Deep Learning such as NLP, image processing

HIDS

- weakness : needs of all configuration files to identify attacks
- solution : allowing access to big data technology

4.1 Proposed Scalable Framework

In order to capture the contextual and sequence related information from system calls, text presentation methods are adopted.

Text Presentation Methods

- Bag - of - Words : Focus on the frequency of words used in system calls.
- N - grams : Focus on the frequency of words and preserve the sequence information of system calls.
- Keras Embedding : Convert the system calls into a numeric according to a lookup table of words.

4.2.1 Proposed Scalable Framework

System calls are transformed into the vector by using the text presentation methods. The vector X becomes an input of DNN.

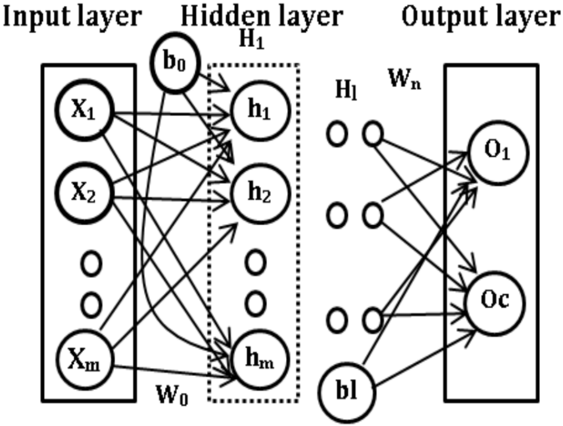


FIGURE 1. Architecture of a deep neural network (DNN).

4.2.2 Proposed Scalable Framework

Hidden layer (H_i) is defined as :

$$h_i = f(w_i^T x + b_i)$$

When it comes to defining h_i as an input to next layer, functions as follows are used :

$$\begin{aligned} sigmoid &= \frac{1}{1+e^{-x}} \\ \tan gent &= \frac{e^{2x}-1}{e^{2x}+1} \\ softmax(x_i) &= \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} \end{aligned}$$

If there many hidden layers an output is as follows :

$$H(x) = H_l(H_{l-1}(H_{l-2}(,,,(H_1(x)))))$$

4.2.3 Proposed Scalable Framework

Loss Function

The prediction loss for Binary classification:

$$loss(pd, ed) = -\frac{1}{N} \times \Sigma_{i=1}^{N}[ed_i logpd_i + (1 - ed_i)log(1 - pd_i)]$$

The prediction loss for Multi - class classification :

$$loss(pd, ed) = -\Sigma_x ed_i logpd(x)log(ed(x))$$

pd : a vector of predicted probablity
ed : a vector of expected class label

5.1.1 Dataset Limitation and Statistical Measures

Term Explanation

Positive : normal access

Negative : attack connection

There are 4 signals : TP (True Positive), TN (True Negative), FP (False Positive), FN (False Negative)

Dataset Limitation

- KDDCup 99 : The oldest and basic dataset
- NSL - KDD : The redundant connection is removed from the KDDCup 99
- UNSW - NB15 : Hybrid of KDDCup and the NSL - KDD
- Kyoto : A honeypot system.
- WSN - DS : A dataset that mainly contains Dos or DDos attack.
- CICIDS 2017 : Most recent dataset.

5.1.2 Dataset Limitation and Statistical Measures

Left figure : t-sne of kddcup 99 dataset
Right figure : t-sne of cicidis2017 dataset

FIGURE 2. t-SNE visualization of KDDCup 99.

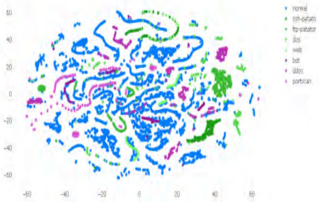


FIGURE 3. t-SNE visualization of CICIDS 2017.

Left Figure : Blue dots are Dos, green dots are normal access.
Right Figure : SKy-blue dots are normal access, right-green dots are brute-force attacks. Pink dots are botnet.

5.2 Dataset Limitation and Statistical Measures

Statistical Measures

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
$$F1 - score = 2(\frac{Precision \times Recall}{Precision + Recall})$$
$$Precision = \frac{TP}{TP + FP}$$
$$TPR = \frac{TP}{TP + FN}$$
$$FPR = \frac{FP}{FP + TN}$$
$$AUC = \int_1^0 \frac{TP}{TP + FN} d\frac{FP}{TN + FP}$$

6.1 Experimental Design

1. Classifying the network connection record as either benign or attack with all features.
2. Classifying the network connection record as either benign or attack and categorizing an attack into its categories with all features.
3. Classifying the network connection record as either benign or attack and categorizing an attack into its categories with minimal features.

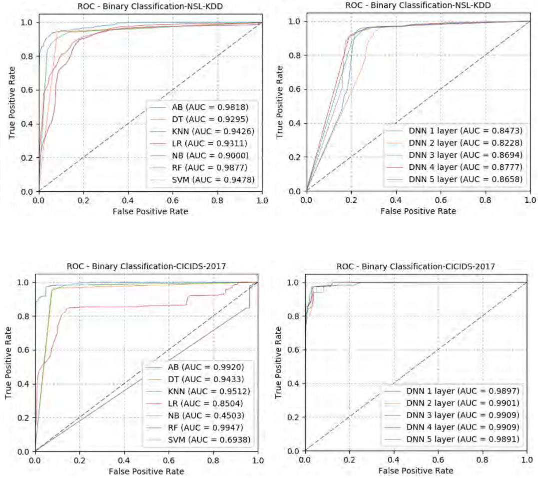
6.2 Experimental Design

Finding Optimal Parameters

Choose the one with the best performance by changing the parameter values

- number of hidden units in DNN : 1024
- learning rate : 0.1
- epochs : 500
- activation functions : ReLU
- dropout rate (rate of removed units in the hidden layers) : 0.01

7.1 Results (ROC, right : DNN, left : others)



In most of the cases, DNN performed better than the classical machine learning classifiers with AUC used as the standard metric.

7.2 Results -Classical algorithms-

All Classifiers

Classical algorithms : LR, NB, KNN, DT, AB, RF, SVM - rbf
DNN : hidden layers are 1, 2, 3, 4, 5

- In terms of accuracy noted that the DT, AB and RF classifiers performed better than the other classifiers
- Additionally, the performance of DT, AB and RF classifiers remains the same.
- $\rightarrow$ This indicates that the DT, AB and RF classifiers are generalizable and can detect new attacks.

7.3 Results -DNNs-

Classical algorithms VS DNNs

In terms of accuracy, the performance of the DNN is clearly superior to that of classical machine learning algorithms.

Text Presentation Methods

Amond Bag-of-Words, N-grams and Keras Embedding, Keras Embedding performed better in terms of accuracy.

Feature Engineering

- Feature selection method significantly reduces the computing time and also showed improved intrusion detection rate.
- The experiments with 11 and 8 feature sets performed well. 11 features performed better.

8 Conclusion and Further Work

Conclusion

- The DNN was chosen by evaluating their performance in comparison to classical classifiers on benchmark IDS datasets.
- We also collected host-based and network-based features in real-time and employed the proposed DNN model, then it succeeded.

Further Work

- The performance of the IDS can be enhanced by adding more configuration such as DNS and BGP events.
- The proposed system does not give detailed information on the structure and characteristics of the malware.
- The execution time of the proposed system can be enhanced by adding more nodes to the existing cluster