DAFL: Deep Adaptive Feature Learning for Network Anomaly Detection

Shujian Ji, Tongzheng Sun, Kejiang Ye, Wenbo Wang, Cheng-Zhong Xu

DAFL: Deep Adaptive Feature Learning for Network Anomaly Detection

Shujian Ji[1,2], Tongzheng Sun[1], Kejiang Ye*[1], Wenbo Wang[3], and Cheng-Zhong Xu[4]

[1] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China
[2] University of Chinese Academy of Sciences, Beijing 100049, China
[3] Khoury College of Computer Sciences, Northeastern University, Seattle WA 98109, USA
[4] Faculty of Science and Technology, University of Macau, Macau, China
{sj.ji, tz.sun, kj.ye}@siat.ac.cn, wang.wenbo@husky.neu.edu, czxu@um.edu.mo

Abstract. With the rapid development of the Internet and the growing complexity of the network topology, network anomaly has become more diverse. In this paper, we propose an algorithm named Deep Adaptive Feature Learning (DAFL) for traffic anomaly detection based on deep learning model. By setting proper feature parameters θ on the neural network structure, DAFL can effectively generate low-dimensional new abstract features. Experimental results show the DAFL algorithm has good adaptability and robustness, which can effectively improve the detection accuracy and significantly reduce the detection time.

Keywords: Network Anomaly Detection · Deep Learning · Feature Learning.

1 Introduction

Network attack is a serious problem in the Internet environment. With the rapid development of the Internet and the growing complexity of the network topology, network anomaly has become more diverse. Network anomaly detection is an effective way to deal with different network attacks [1].

Machine learning is a common method for anomaly detection in the network environment, such as Naive Bayes, Support Vector Machine and other shallow learning technologies [2,3]. Although these technologies have improved the detection accuracy to a certain extent, they also face some limitations. For example, expert knowledge is required for data processing, and a large amount of time is needed for data training. Recently, deep learning based methods [4–6] are proposed for anomaly detection due to the better feature learning ability. However, they improved the detection accuracy, without taking into account the training time and execution time in high-speed networks.

* Corresponding author

In this paper, we propose an algorithm named Deep Adaptive Feature Learning (DAFL) which can utilize the feature learning ability of deep learning and the advantages of transfer learning. The contributions of this paper are summarized as follows: (i) the algorithm can determine the structure of the neural network according to the dimension of data. (ii) By combining deep learning with shallow machine learning, DAFL improves the classification performance of anomaly detection and greatly reduces the training time.

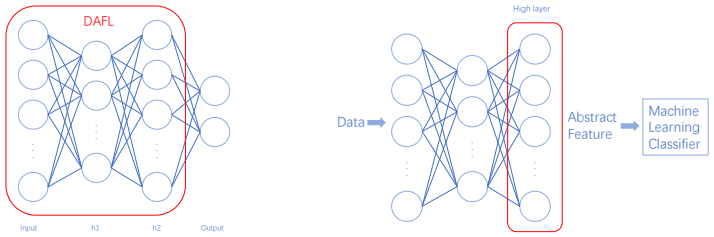
2 DAFL Algorithm

we design the DAFL algorithm to determine the number of layers of the network hidden layer and the number of neurons in each layer according to the dimensions of the input data, and construct a pre-trained learning model that can adapt to the dimension of data features, as shown in Algorithm 1.

Algorithm 1 Deep Adaptive Feature Learning

Require: training sample v, feature parameter θ, learning rate η, list N
Ensure: pre-train model $xW + b$
D = data dimension of v, layer number: $\mathrm{l} = \lceil D/5 \rceil$
inital $n_1 = D$, calculate neurons number of each layer: $n_l = \lceil \theta * D \rceil$
for $i = 2$ to $l - 1$ **do**
    neurons: $n_i = \lceil D/i^2 \rceil + \lceil \theta * D \rceil$, save n_i to N
end for
for $i = 1$ to l **do**
    use N to build network with l-th layers and n_i neurons.
    output layer: $S(x) = \frac{1}{1+e^{-x}}$
end for
for all v_i **do**
    calculate the actual output of the neuron $v_i^{'}$
    $\delta_k = v_i^{'}(1 - v_i^{'})(v_i - v_i^{'})$
    hidden layer h error gradient: $\delta_h = v_h^{'}(1 - v_h^{'})W_{hk}\delta_k$
    update the weights: $W_{ij} = W_{ij} + \Delta W_{ij}$, $\Delta W_{ij} = \eta O_i \delta_j$, update the bias: $b_j = b_j + \eta\delta_j$
end for

In order to balance the training speed and accuracy of the deep learning model, we design a feature parameter θ (from 0.1 to 1) as the control value in the DAFL algorithm to make the high hidden layer generate abstract features of different dimensions. As shown in Fig. 1, combining the deep network structure based on DAFL with different conventional shallow machine learning classifiers can be used as the detection model.



(a) Pre-train DBN model by using DAFL

(b) Use the remain structure to generate new feature

Fig. 1. The design of DAFL

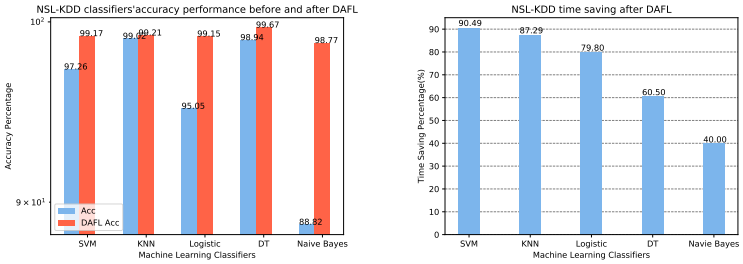
3 Experiment

We conduct experiments with NSL-KDD [7] dataset to evaluate our proposed algorithm. By comparing the performance of the original data and the data processed by the DAFL algorithm on the classifier, we can verify the validity of DAFL.

Table 1. Models Performance in NSL-KDD Dataset

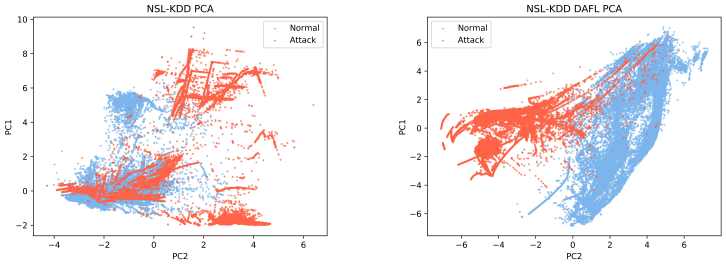
Model	Accuracy	Precision	Recall	$F_1 - score$	Time(s)
Support Vector Machine(SVM)	97.26%	98.03%	96.06%	97.19%	91.73s
DAFL SVM(θ=0.8)	**99.17**%	**99.49**%	**98.74**%	**99.15**%	**8.72*s***
K-Nearest Neighbors(KNN)	99.02%	99.34%	98.88%	99.15%	107.19s
DAFL KNN (θ=0.8)	**99.21**%	**99.41**%	**98.89**%	**99.19**%	**13.62*s***
Logistic Regression(LR)	95.05%	95.26%	94.05%	94.98%	1.98s
DAFL LR(θ=0.8)	**99.15**%	**99.36**%	**98.82**%	**99.13**%	**0.40*s***
Decision Tree(DT)	98.94%	98.98%	98.75%	98.93%	1.19s
DAFL DT(θ=0.8)	**99.67**%	**99.72**%	**99.58**%	**99.65**%	**0.47*s***
Naive Bayes(NB)	88.82%	86.72%	89.24%	88.66%	0.10s
DAFL NB(θ=0.8)	**98.77**%	**95.25**%	**99.75**%	**97.70**%	**0.06*s***

Experiments show that the classifier achieves the best result when the feature parameter is set to 0.8. Table 1 shows the changes of classifier performance metrics when DAFL is applied to the classifier on the NSL-KDD dataset. It is worth noting that the accuracy in the NB classifier increased from 88.92% to 98.77%, and the recall increased from 89.24% to 99.75%. In terms of detection time, the classifier that has been processed by the DAFL algorithm has a significant reduction in detection time. The most obvious change is that the time of the SVM classifier is reduced from 91.73s to 8.72s, and the detection time of the KNN classifier is reduced from 107.19s to 13.62s. Fig. 2 shows the accuracy and time saving percentage on NSL-KDD.



(a) Classifiers' accuracy performance comparison on NSL-KDD (b) NSL-KDD time saving after DAFL

Fig. 2. Accuracy and time saving on NSL-KDD



(a) PCA visualization for NSL-KDD (b) PCA visualization for NSL-KDD after DAFL

Fig. 3. PCA visualization for NSL-KDD before and after DAFL

We performe data scatter visualization by PCA method for normal traffic and abnormal traffic in the dataset in Fig. 3. It is obvious that the DAFL algorithm can separate the normal traffic and abnormal traffic.

4 Related work

There are a lot of work on network anomaly detection. Ibrahimi *et al.* used classification algorithms such as linear discriminant analysis (LDA) and principal component analysis (PCA) to classify abnormal network traffic [8]. Alrawashdeh *et al.* used Restricted Boltzmann Machine (RBM) to perform unsupervised feature reduction [9]. Potluri *et al.* proposed an accelerated DNN structure for identifying network data anomalies [10]. Kang *et al.* proposed an intrusion detection system based on deep neural network [11]. Our research group also proposed different algorithms and tool for network anomaly detection [1, 4–6].

5 Conclusion

In this paper, we propose a DAFL algorithm for network anomaly detection that can determine the number of hidden layers and the number of neurons in each hidden layer according to the dimension of the original data. Using the idea of transfer learning, we remove the output layer of the neural network and use the residual structure to generate new data with abstract features as input of other machine learning classifiers. The experimental results show that the method achieves good results, and has a certain degree of robustness and adaptability.

Acknowledgment. This work is supported by the National Key R&D Program of China (No. 2018YFB1004804), National Natural Science Foundation of China (No. 61702492), Shenzhen Discipline Construction Project for Urban Computing and Data Intelligence, and Shenzhen Basic Research Program (No. JCYJ20170818153016513).

References

1. P. Lin, K. Ye, and C.-Z. Xu, "Netdetector: an anomaly detection platform for networked systems," in *IEEE International Conference on Real-time Computing and Robotics*. IEEE, 2019.
2. T. Shon, Y. Kim, C. Lee, and J. Moon, "A machine learning framework for network anomaly detection using svm and ga," in *Proceedings from the sixth annual IEEE SMC information assurance workshop*. IEEE, 2005, pp. 176–183.
3. N. B. Amor, S. Benferhat, and Z. Elouedi, "Naive bayes vs decision trees in intrusion detection systems," in *Proceedings of the 2004 ACM symposium on Applied computing*. ACM, 2004, pp. 420–424.
4. P. Lin, K. Ye, and C.-Z. Xu, "Dynamic network anomaly detection system by using deep learning techniques," in *International Conference on Cloud Computing*. Springer, 2019, pp. 161–176.
5. M. Zhu, K. Ye, Y. Wang, and C.-Z. Xu, "A deep learning approach for network anomaly detection based on amf-lstm," in *IFIP International Conference on Network and Parallel Computing*. Springer, 2018, pp. 137–141.
6. M. Zhu, K. Ye, and C.-Z. Xu, "Network anomaly detection and identification based on deep learning methods," in *International Conference on Cloud Computing*. Springer, 2018, pp. 219–234.
7. "Nsl-kdd," *"https://iscxdownloads.cs.unb.ca/iscxdownloads/NSL-KDD"*, 1999.
8. K. Ibrahimi and M. Ouaddane, "Management of intrusion detection systems based-kdd99: Analysis with lda and pca," in *2017 International Conference on Wireless Networks and Mobile Communications (WINCOM)*. IEEE, 2017, pp. 1–6.
9. K. Alrawashdeh and C. Purdy, "Toward an online anomaly intrusion detection system based on deep learning," in *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, 2016, pp. 195–200.
10. S. Potluri and C. Diedrich, "Accelerated deep neural networks for enhanced intrusion detection system," in *2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA)*. IEEE, 2016, pp. 1–8.
11. M.-J. Kang and J.-W. Kang, "Intrusion detection system using deep neural network for in-vehicle network security," *PloS one*, vol. 11, no. 6, p. e0155781, 2016.