A Novel Intrusion Detection Method for Internet of Things

Peisong Li, Ying Zhang*

College of Information Engineering, Shanghai Maritime University, Shanghai 201306
E-mail: yingzhang@shmtu.edu.cn

Abstract: Internet of Things (IoT) era has gradually entered our life, with the rapid development of communication and embedded system, IoT technology has been widely used in many fields. Therefore, to maintain the security of the IoT system is becoming a priority of the successful deployment of IoT networks. This paper presents an intrusion detection model based on improved Deep Belief Network (DBN). Through multiple iterations of the genetic algorithm (GA), the optimal network structure is generated adaptively, so that the intrusion detection model based on DBN achieves a high detection rate. Finally, the KDDCUP data set was used to simulate and evaluate the model. Experimental results show that the improved intrusion detection model can effectively improve the detection rate of intrusion attacks.

Key Words: Internet of Things, Intrusion detection, Deep Belief Network

1 INTRODUCTION

With the rapid development, IoT technology has been widely used. However, IoT has become an ideal target by cyber attacks because of its distributed nature, number of objects and openness [1-5]. In addition, because many IoT nodes collect, store and process private information, they come to become the apparent target by malicious attackers [6].

There are many intrusion detection models, such as statistical analysis based [7], cluster analysis based [8], artificial neural network [9] or deep learning based [10]. Among these methods, intrusion detection based on deep learning has demonstrated a better performance than the traditional methods [11]. Previous research shows that the accuracy of the intrusion detection model based on deep learning method is affected by the number of hidden layers and the number of neurons in each layer. Inappropriate network structure will have a relatively large impact on detection rate. In the past few years, there has not been a unified solution for the selection of the number hidden layer and the number of neurons. Most of the research is based on trial and error and on pruning or constructive methods [12], the network structure and the performance cannot be guaranteed.

This paper proposes an intrusion detection model based on improved Deep Belief Network (DBN) and Genetic Algorithm (GA). For different attacks including low-frequency attacks and other types of attacks, the corresponding different optimized network structures of DBN are obtained by iterative evolution, and the DBN with obtained most optimal network structure will be used for intrusion detection. By applying the GA, iteratively generates the optimal number of hidden layers and neurons in a hidden layer, and reduces the network complexity as much as possible while ensuring the detection rate. This method can make the intrusion detection system have a higher detection rate and greater improvement in performance.

This paper will firstly introduce the related work of intrusion detection based on machine learning. Then introduce the proposed algorithm model. In the end, the experimental results and compared it with other methods will be shown.

2 RELATED WORK

Intrusion detection technology based on artificial neural network is generally divided into three categories: intrusion detection technology based on supervised artificial neural network, unsupervised intrusion detection technology and intrusion detection technology based on hybrid neural network.

The main advantage of the unsupervised artificial neural network is that new data can be analyzed without tagging data in advance. The Self-Organizing Feature Map (SOM) used in [13] is an unsupervised learning method that extracts features from normal system activity and identifies statistical changes from normal trends. However, for low-frequency attacks, the detection accuracy of unsupervised neural network is also low.

Supervised neural networks mainly include multilayer feed-forward (MLFF) neural networks. Singh et al. [14] use MLFF Neural Networks based on user behavior to detect anomaly. However, sometimes the distribution of training data sets is not balanced, which makes the MLFF neural network easily reach the local minimum value, and thus the stability is low.

The third category is the hybrid neural network model, and such a model, FC-ANN is proposed in [15]. The FC-ANN method introduces fuzzy clustering techniques into general artificial neural networks. Salama et al. [16]

This work is supported by National Natural Science Foundation of China under Grant 61673259.

*: The correspondent author.

978-1-7281-0106-4/19/$31.00 ©2019 IEEE

proposed an intrusion detection model combining deep belief network (DBN) and support vector machine (SVM). Firstly, the dimensionality of feature set is reduced by DBN, and then SVM is used for classification.

At present, there are many intrusion detection technologies based on deep learning. Abolhasanzadeh [17] proposed a method for detecting attacks in big data using Deep Auto-Encoder. Gao et al. [18] trained the deep belief network (DBN) as a classifier to detect intrusions.

3 THE PROPOSED ALGORITHM MODEL

3.1 DBN for intrusion detection

DBN is composed of multiple Restricted Boltzmann Machines (RBMs), mainly executing unsupervised learning of pre-processed data, processing and abstracting high-dimensional data [19]. DBN module is mainly divided into two steps in the training model:

(1) Each RBM is trained separately, characterized by unsupervised and independent.

The observed joint distribution of the input value x and hidden layer H_k is modeled as follows:

$$P(x, H_1, \dots, H_N) = (\prod_{k=0}^{N-2} P(H_k \mid H_{k+1})) \bullet P(H_{N-1}, H_N) \qquad (1)$$

where $x = H_0$, $P(H_{k-1} \mid H_k)$ is a conditional distribution of visible units in the k layer with the condition of hidden units of RBM. $P(H_{N-1}, H_N)$ is the visible-hidden joint distribution at the top level of the RBM. The illustration is as follows:

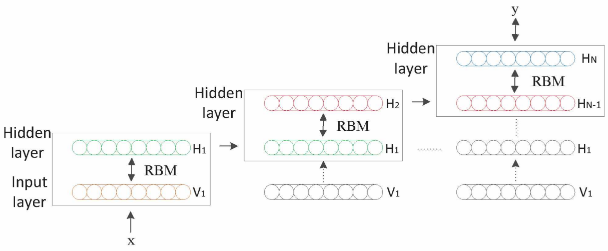

Fig. 1 RBM training process

The first layer is trained as an RBM, assigning the x input to V_1 as the visible layer.

The input data obtained from the first layer is characterized as the second layer's data.

The second layer is trained as an RBM, and the transformed data is used as a training sample.

Finally, repeat this process until to the last layer. So that it is a Deep Learning method.

(2) The last layer of the DBN is the BP neural network. According to the characteristics of the BP neural network, the BP neural network can propagate error information from top to bottom in each layer of RBM, fine-tune the DBN network, and achieve global optimization.

The number of hidden layers and the number of neurons in each layer in the deep belief network are determined by the algorithm model constructed earlier.

3.2 Multiple iterations

GA is known to be an ideal technique for finding optimal solutions to various problems.

1） Population initialization

Initialization operation is to generate an initial population randomly for subsequent genetic manipulation. For a simple training set, up to three hidden layers are enough to get a good detection rate. The number of nodes in the three hidden layers is encoded directly in the binary chromosome. The length of chromosome is 18 bits: the first 6 bits are reserved for the first hidden layer, the subsequent 7-12 bits and 13-18 bits are for the second and the third hidden layers respectively, as shown in Fig. 2:

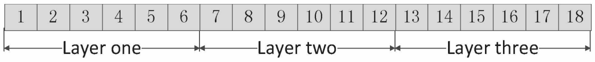

Fig. 2 Chromosome schematic

A chromosome represents a network structure, which has at most three hidden layers and at least one hidden layer. When the population is initialized, the number of nodes in each layer is smaller than the number of input features and greater than the number of output features must be ensured.

$$I \le N \le O \qquad (2)$$

where I is the size of the input layer, O is the size of the output layer and N is the number of neurons in the hidden layer.

2） Selection

Selection operation is to select excellent chromosomes from the current population and prepare for the following operations. In general, a method of roulette wheel selection based on proportional fitness assignment is used. As the fitness of candidate individuals increases, the probability of being selected increases. So, the individuals will be selected according to the method of roulette. This ensures that the best individuals will not be eliminated.

3） Crossover

Crossover operation using partially matched crossover (PMC).

To use this exchange method can avoid falling into a local optimum, thus the diversity of the next generation can be increased and the convergence rate can be accelerated. At the same time, another possibility is the number of hidden layers of the intersecting individuals is different. For this case, the method of randomly selecting a layer common to both chromosomes to crossover will be adopted. This is done to avoid the situation that the number of neurons in an intermediate hidden layer is 0. Method is demonstrated in Fig. 3:

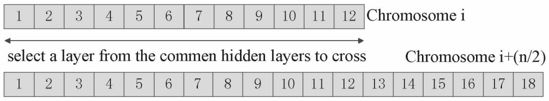

Fig. 3 Crossing chromosomes with different hidden layers

4） Mutation

Mutation operation is to change a certain bit in the chromosome. It can use the random search ability of mutation operator. When the operation result is close to the optimal solution neighborhood, it can quickly converge to the optimal solution.

3.3 Algorithm flow

The algorithm flow chart is as follows:

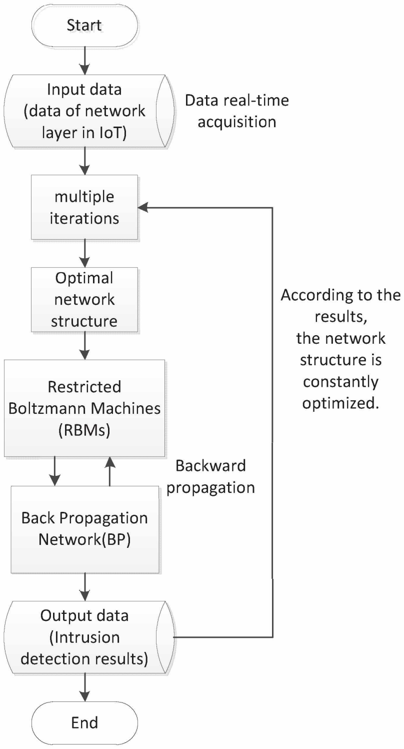

Fig. 4 Algorithm model flow chart

4 EXPERIMENTAL SIMULATION

4.1 Experimental data

The KDDCUP data set was generated by an intrusion detection assessment project of the US Department of Defense Advanced Planning Agency (DARPA) [20]. This data set is now used for network intrusion detection.

In addition, the KDDCUP data set needs to be normalized. The method used in this paper is the Min-Max normalization method, which mapping the resulting value to [0, 1], the conversion function is as follows:

$$X^* = \frac{X - Min}{Max - Min} \tag{3}$$

where Max is the maximum value of the sample data, and Min is the minimum value of the sample data.

4.2 SIMULATION RESULTS

The DoS, R2L, Probe, U2R four classes of attacks are selected as intrusion attack training sets respectively [21-22]. The optimal chromosome generated by the iteration is decoded, and then the optimal network structure is obtained as shown in Table 1:

Table 1. Optimal network structure for different types of attacks

Number	Attack	Network Structure
A	DoS	41-18-12-2
B	R2L	41-31-2
C	Probe	41-26-2
D	U2R	41-38-2

The network structure of the DBN includes the input layer, hidden layer, and output layer. The size of input layer is 41 and output layer is 2. The middle is hidden layer. For data sets with different attack, the different optimal network structure is generated by multiple iterations of the GA.

Intrusion detection is performed on four classes of attacks using the A-D network structures respectively, and their detection rates are calculated. Shown in table 2:

Table 2. Detection rate for different class of attack

Structure	DoS	R2L	Probe	U2R
type A	**99.45%**	95.18%	90.33%	97.27%
type B	97.60%	**97.78%**	98.23%	96.38%
type C	70.00%	95.02%	**99.37%**	98.27%
type D	61.23%	86.32%	98.35%	**98.68%**

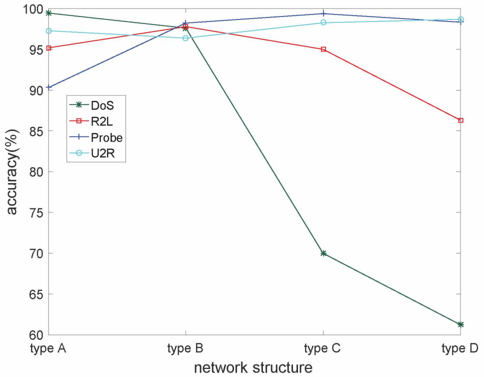

Fig. 5 Detection rate for different class of attack

As seen from Fig. 5, for a certain type of network structure generated by the certain type of attack, the detection rate of this type of network is higher than other network structures.

Meanwhile, this paper compared our method with the methods TANN, FC-ANN, SA-DT-SVMS, and BPNN proposed by others. The results obtained are compared with the above methods and summarized in the following table:

Table 3. Classification accuracy of each method

Method	DoS	R2L	Probe	U2R
FC-ANN	96.70%	93.18%	48.12%	83.33%
TANN	90.94%	80.53%	94.89%	60.00%
SA-DT-SVMS	100.00%	93.22%	98.36%	80%

BPNN	80.35%	89.12%	89.12%	25.58%
GA-DBN	**99.45%**	**97.78%**	**99.37%**	**98.68%**

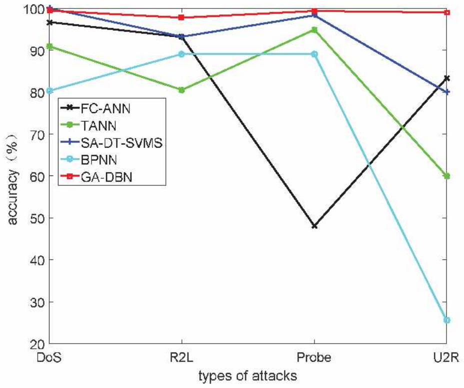

Fig. 6 Classification accuracy of each method

As seen from Fig. 6 that the proposed GA-DBN method has reached a very high level for the detection of four types of attacks. The classification accuracy of DoS is higher than 99%, and the classification accuracy of R2L, Probe and U2R is also significantly higher than other methods.

5 CONCLUSION

In this paper, GA performs multiple iterations to produce an optimal network structure, and DBN then uses the obtained network structure as an intrusion detection model to classify, in this way solved the problem of how to select an appropriate network structure when using deep learning methods for intrusion detection, and thus improve the classification accuracy.

This method has many advantages: on the one hand, the specific network structure generated for specific attack types is higher in classification accuracy than other network structures, which can reach more than 99%. On the other hand, for small training sets, such as U2R, the classification accuracy of our algorithm is also significantly higher than other methods.

In the future, the optimization of other parameters of the deep network will be considered.

REFERENCES

[1] BB Zarpelao, RS Miani, and CT Kawakani, “A survey of intrusion detection in Internet of Things,” Journal of Network and Computer Applications, vol. 84, pp. 25-37, 2017.

[2] A. Abduvaliyev, A.K Pathan, J. Zhou, R. Roman and W. Wong, “On the Vital Areas of Intrusion Detection Systems in Wireless Sensor Networks”, Communications Surveys & Tutorials, IEEE vol. 15, pp. 1223-1237, 2013.

[3] K Yang, J Ren, Y Zhu, and W Zhang, “Active Learning for Wireless IoT Intrusion Detection,” IEEE Wireless Communications, vol. 25, no. 6 pp. 19-25, 2018.

[4] P Schulz, M Matthe, H Klessig, M Simsek, et al. “Latency critical IoT applications in 5G: Perspective on the design of radio interface and network architecture,” IEEE Communications Magazine vol. 55, no. 2, pp. 70-78, 2017.

[5] A Ahmad, MH Rehmani, H Tembine, et al. “IEEEAccessSpecial Section Editorial: Optimization for Emerging Wireless Networks: IoT, 5G, and Smart Grid Communication Networks,” IEEE Access vol. 5, pp. 2096-2100, 2017.

[6] H HaddadPajouh, A Dehghantanha, R Khayami, KK Choo, “A deep Recurrent Neural Network based approach for Internet of Things malware threat hunting”, Future Generation Computer Systems, vol. 85, pp. 88–96, 2018.

[7] AA Kuznetsov, and AA Smirnov, “The statistical analysis of a network traffic for the intrusion detection and prevention systems," Telecommunications and Radio Engineering, vol. 74, no. 1, 2015.

[8] N Pandeeswari, and G Kumar, “Anomaly detection system in cloud environment using fuzzy clustering based ANN,” Mobile Networks and Applications, vol. 21, no. 3 pp. 494-505, 2016.

[9] SS Roy, A Mallik, R Gulati, and MS Obaidat, “A deep learning based artificial neural network approach for intrusion detection,” In International Conference on Mathematics and Computing, pp. 44-53, Springer, Singapore, 2017.

[10] A Javaid, Q Niyaz, and W Sun, “A deep learning approach for network intrusion detection system,” In Proceedings of the 9th EAI International Conference on Bio-inspired Information and Communications Technologies (formerly BIONETICS), pp. 21-26. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2016.

[11] R. Beghdad, “Critical study of neural networks in detecting intrusions,” Computers and Security, vol. 27, pp. 168-175, 2008.

[12] S. Mukkamala, G. Janoski, and A. Sung. “Intrusion detection using neural networks and support vector machines,” Proceedings of the International Joint Conference on Neural Networks (IJCNN’02), Honolulu, HI, USA, pp. 1702–1707, 2002.

[13] A Saraswati, M Hagenbuchner, Zhi Quan Zhou, “High Resolution SOM Approach to Improving Anomaly Detection in Intrusion Detection Systems”, Carcinogenesis, vol.9992, pp.947-954, 2016.

[14] N Singh, and A Kaur, “A survey: Multilayer feed-forward neural network approaches for intrusion detection system,” International Journal For Technological Research In Engineering, vol. 2, no. 11, pp. 2906-2907, 2015.

[15] G. Wang, J. Hao, J. Ma, L. Huang, “A new approach to intrusion detection using artificial neural networks and fuzzy clustering”, Exp. Syst. Appl. pp. 6225–6232, 2010.

[16] M. Salama, H. Eid, R. Ramadan, A. Darwish, and A. Hassanien, “Hybrid intelligent intrusion detection scheme,” in Soft Computing in Industrial Applications, vol. 96. Berlin, Germany: Springer-Verlag, ser. Advances in Intelligent and Soft Computing, pp. 293–303, 2011.

[17] B. Abolhasanzadeh, “Nonlinear dimensionality reduction for intrusion detection using auto-encoder bottleneck features,” in 2015 7th Conference on Information and Knowledge Technology (IKT), Urmia, Iran, pp. 1–5, 2015.

[18] N. Gao, L. Gao, Q. Gao, and H. Wang, “An Intrusion Detection Model Based on Deep Belief Networks,” in 2014 Second International Conference on Advanced Cloud and Big Data, Huangshan, China, pp. 247–252, 2014.

[19] G. E. Hinton, S. Osindero, and Y. Teh, “A fast learning algorithm for deep belief nets,” Neural Computation, vol. 18, pp. 1527–1554, 2006.

[20] V Sandulescu, and M Chiru, “Predicting the future relevance of research institutions-The winning solution of the KDD Cup 2016,” arXiv preprint arXiv: 1609.02728, 2016.

[21] MJ Rausch, VB Krishna, P Gu, et al. “Peer-to-peer Detection of DoS Attacks on City-Scale IoT Mesh Networks,” In 2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), pp. 1-6. IEEE, 2018.

[22] AA Diro, and N Chilamkurti, “Distributed attack detection scheme using deep learning approach for Internet of Things,” Future Generation Computer Systems vol. 82, pp: 761-768, 2018.