2528-9705

Örgütsel Davranış Araştırmaları Dergisi
Journal Of Organizational Behavior Research
Cilt / Vol.: 4, Sayı / Is.: S2, Yıl/Year: 2019, Kod/ID: 91S2489

INTRUSION DETECTION IN COMPUTER NETWORKS THROUGH A HYBRID APPROACH

Mehdi KHODAMORADI

Department of Computer Engineering, Faculty of Technical and Engineering, University of Imam Reza International, Mashhad, Iran.

ABSTRACT

The existence of intrusion detection systems is important, because despite all security mechanisms, a system has many vulnerable points which increase the possibility of attack from these system. Therefore, in this research, Intrusion Detection in Computer Networks is considered with a hybrid approach. The proposed method is a three-step approach that attempts to improve the attacks detection and reduce the false alarms. First, the multi-class problem has been converted into several two-class problems and then, appropriate properties of each class is extracted based on a various approaches of information gain and Fisher algorithm. In the last step, we will have one output for each classifier and four outputs for each class. Since the number of dataset classes are examined by five ordinary classes, denial of service attack, port scanning attack, remote local access attack and user attack to the root, and 20 output results of different classifiers are created for each sample. The used data set in this research is KDD-CUP99. The most important used evaluation criteria include precision, readout, false alarm rate, F-criterion, and error rate. In the proposed algorithm, decision tree, naive Bayes, K- nearest neighbor and the neural network have been used as an initial classifier and an incremental algorithm based on the decision tree has been used as the final classifier. The proposed method in all classes could be able to perform better function than the previous method.

Keywords: *Intrusion Detection System, Proposed Approach, Spatial Mapping Method.*

INTRODUCTION

Any unauthorized access or an activity against an Informative/Communicative system either inadvertently or deliberately is called intrusions (Coolen and Luiijf, 2002). Systems which detect destructive and intrusive factors are called Intrusion Detection Systems (IDS). These sort of systems, by evaluating and analyzing the exchanged data and current traffic, try to detect the destructive factors and after detecting and making sure that the related traffic relates to an unauthorized machine or user, it informs the network administrator in various ways. Intrusion Detection Systems are used in two layers of network and application. The existence of Intrusion Detection Systems is important because despite all security mechanisms including verification, identification, encryption and decryption, firewall, access control, etc., system still has high vulnerable points (Catania and Garino, 2012).

Some research has been done in this field (Kaplantzis and Mani, 2006) and considered classifier techniques applied in attack detection. In this research, clustering methods including K-means, neural networks and support vector machines are used in attack detection and the research results show the highest precision degree among the selection tools for support vector machines (Kang, Fuller and Honavar, 2005).Using a part of systematic calling of host operating system and polynomial naive Bayes classifier systems, decision tree, C4, support

Geliş tarihi/Recieved: 26.11.2018 – Kabul tarihi/Accepted: 24.03.2019 – Yayın tarihi/Published: 21.05.2019

vector machine and logistic regression could achieve a high degree of precision in attack detection.

In (Horng et al., 2011), the support vector machine is used as a classifier method to solve the intrusion detection problem. But since this problem is not a dual classifier problem (the attacks detection makes the number of classes be more than 2) so there must be several separate support vector machines used. Research like (Hofman, Schmitz and Sick, 2003; Kayacik, Zincir-Heywood and Heywood, 2003) has used neural networks based on radial function for learning multiple local clusters to detect distinctive attacks and normal events.

The decision trees have also many applications in Intrusion Detection Systems. For example, reference (Depren et al., 2008) has considered the intrusion detection in two sections of known attacks and unknown attacks. Based on the research, the Bayes networks also due to features extraction ability have important applications in the Intrusion Detection Systems (Langely, Iba and Thomas, 1992). The reference (Muda et al., 2011) has used the combination of the K-Mean clustering and naive-Bayes to improve the Intrusion Detection System. The reference (Xu, An and Geng, 2011) has used the combination of Kernel Principal Component Analysis (KPCA) methods, Particle swarm optimization (PSO) and basic radial functions to improve the classifier in intrusion detection. Today, the Intrusion Detection Systems use the machine learning methods and techniques. So far, many different methods and solutions using the machine learning techniques have been provided for this purpose. The use of hybrid or ensemble methods have been grown in recent years (Shelly Xiaonan and Banzhaf, 2010; Catania and Garino, 2012), so improving the accurate intrusion detection rate in these kind of systems is so important.

The extent application of hybrid methods in recent years shows that the use of these methods leads to better results. Considering the use of hybrid methods give many capabilities for improving the precision of detecting the type of attacks in computer networks.

THE PROPOSED METHOD

The proposed structure is three-step and in the initial step, preprocessing is performed on data and then the obtained data is assigned to classifiers in order to create mean results and after creating mean results, they are sent to the final step for training. Initially, in order to choose feature using different approaches like Fisher algorithm and interest information based on calculation entropy, the usefulness of each feature is extracted. Then with the help of various clustering method and the training samples, they are chosen and given to the initial classifiers . In the following, the initial classifiers are made and sent to the next step. So, we will have two outputs for each classifier and according to the number of used initial classifiers in the mean results, step in output data for each sample from the second part in the training set (n is the number of classifiers of mean results part).Finally using the obtained results the mean data is formed and sent to classifier of the final layer for training.

Feature Selection component

To evaluate features and choose the most valuable of them, two approaches of Fisher algorithm and interest information were used. In the proposed method, the top 10 features of the Fisher method are used in neural network classifier. Equation (1) shows the Fisher rate.

$$FR(x) = \frac{(\bar{x}_1 - \bar{x}_2)}{\sigma_{x_1}^2 + \sigma_{x_2}^2} \qquad (1)$$

According to equation (1), parameter x is feature values and x is mean and and σ variance of samples belonging to each class. In the proposed method, the top 10 features obtained from method of Fischer are used to classify K, the nearest neighbor. In equations (2) to (3-4) the way of calculating and extracting is considered based on

$$S=Info(x) - Info_x(x) \qquad (2)$$

$$Info(X) = -\sum_{i=1}^{k} p(c_i, X) \times \log(p(c_i, X)) \qquad (3)$$

$$Info_x(x) = -\sum_{I=1}^{V} \frac{|V_i|}{|X|} \times Info(V_i) \qquad (4)$$

The value of information usefulness of each features is obtained according to the equation (3). The amount of P (c_i, X) is equal to the amount of probability density function of data of a unique feature with the same label. Parameter K =2 shows the number of classes and V shows the number of unique values. Infox (X) shows the entropy of classes and $Info_x$(X) shows the entropy of class based on feature value x (Chun et al., 2014).

Sample selection component

In this research, using sample deleting methods and smarter selection of samples try to balance the data set. In order to increase diversity and also increase the accuracy of the various classifiers 4 methods of random selection, random selection based on the K-mean clustering, random selection have been used based on self-organized mapping clustering and random selection based on distance to data average in each class.

Data Normalization

One of the most important ways of normalization is converting data to a new set in which all values from the current range are converted to a new interval and range.
For this purpose, we can use the equation (5):

$$V = \frac{V - min_A}{max_A - min_A}(new\_max_A - new\_min_A) + new\_min_A \qquad (5)$$

That min_A and max_A show in order minimum and maximum values of columns containing feature values, new-min_A shows the new minimum value and new-min_A shows the new maximum value.

Primary classifier component and Creating mean results

In this research, the studied data has 5 different classes: 1. Denial of Service Attacks 2. Probing 3. Remote to User Attacks 4. User to Root Attacks and standard behavior. Therefore, the Output for this component to create mean values would include twenty values. After producing these twenty values for the whole samples of the tutorial sets, the correct tag of the samples will be

added in a new column. In fact, the obtained matrix form the mean dataset that uses for training in the step after this set. In order to create mean results in the proposed method, four different classifiers has been used: decision tree, the closest neighbor K, simple Bayesian and the neural network.

EVALUATION AND RESULTS

In order to evaluate the proposed method, Tutorial Dataset and KDD-CUP 99 test has been used. The obtained results of the implementation of the different parts of the proposed algorithm include the preprocessing step, creating mean results and the use of the final class and the influences of each step will be presented. Finally a comparison is done between the proposed method and other previous methods and the reasons for result improvement in the proposed hybrid approach has been made. The evaluation of the performance of the proposed intrusion detection system would be done in the following sections.

Performance Evaluation and the effects

of sample selection on clustering method of the four random selection methods, intervals based on distance to class mean, clustering based on K-Mean clustering method and clustering based on Self-organizing Network has been used in order to select samples for learning the proposed system.

Table 1: Final results of the proposed method

Proposed Approach	Random		
0.736468	0.718107	Precision	Normal class
0.994818	0.992474	Recall	
0.846367	0.833287	F-Value	
0.998887	0.997753	Precision	DOS class
0.972274	0.964142	Recall	
0.985401	0.98066	F-Value	
0.849577	0.834403	Precision	Prob Class
0.820211	0.74868	Recall	
0.834636	0.789221	F-Value	
0.808023	0.830203	Precision	R2L class
0.069003	0.067597	Recall	
0.127149	0.125014	F-Value	
0.529412	0.168224	Precision	U2R Class
0.257143	0.257143	Recall	
0.346154	0.20339	F-Value	

As shown in table 1, sample selection with various implied methods in the proposed method increased F criteria in standard class, Probing and User to Root Attacks. In two classes of Denial of Service Attacks and Remote to User Attacks, the results are more close to each other compared to the F-criteria.

Evaluation of the performance of the proposed method along with subsystems

Table 2: Final results of the the proposed method

Correct	Actual	U2R	R2L	Prob	DOS	Normal	Predict/Real
0.994818	60591	8	29	217	60	60277	Normal
0.972274	229853	0	227	368	223480	5778	DOS
0.820211	4166	0	2	3417	187	560	Prob
0.069003	16347	8	1128	20	2	15189	R2L
0.257143	70	18	10	0	0	42	U2R
	311027	34	1396	4022	223729	81846	Actual
		0.529412	0.808023	0.849577	0.998887	0.736468	Correct

As shown in table 2, The Denial of Service Attacks detection declaration rate is 97.22 Percent in correct cases and 2.77 percent in false cases and also Denial of Service Attacks in correct cases is equal to 99.88 percent. In the case of probing class samples, detection declaration rate in correct cases is 82.02 percent with detection precision rate of 84.95 percent Also 15.04 percent of other classes samples were put in probing class by mistake.

Remote to User Attacks detection declaration rate in correct cases is 6.90 percent and the detection precision is 80.80 percent. Also 93.3 percent of samples were put in other classes by mistake. Precision of detection in samples of User to Root Attacks is 52.94 percent, detection declaration rate in correct cases is 25.71 Percent and in cases of false declaration is 74.29 percent.

comparing other classes instead of final classifier of the proposed method

Table 3: examining other classifiers as the final classifier

Tree	*KNN*	*NN*[1]	*Adaboost*		
0.907861	0.908488.	0.908739	0.913874	*Recall*	*Normal*
0.997906	0.997532	0.998464	0.99863	*Percision*	
0.950756	0.95093	0.951491	0.954374	*F-Value*	
0.972613	0.972104	0.972917	0.972274	*Recall*	*DOS*
0.99619	0.997651	0.995672	0.998887	*Percision*	
0.98426	0.984712	0.984163	0.985401	*F-Value*	
0.655545	0.785406	0.710514	0.820211	*Recall*	*Prob*
0.9324	0.912946	0.912454	0.849577	*Percision*	
0.769838	0.844387	0.79892	0.834636	*F-Value*	
0.016211	0.02661	0.004588	0.069003	*Recall*	*R2L*
0.543033	0.960265	0.925926	0.808023	*Percision*	
0.031482	0.051786	0.009131	0.127149	*F-Value*	
0.057143	0.185714	0.042857	0.257143	*Recall*	*U2R*
0.5	0.168831	0.5	0.529412	*Percision*	
0.102564	0.176871	0.078947	0.346154	*F-Value*	

[1] Neural Network

By examining Table 3, we can claim that in the standard class and Denial of Service Attacks All classes have almost the same function and there is not much distinction. In the probing class, the best performance is related to increasing classifier Based on decision tree and then K the closest neighbor. In Remote to User Attacks and User to Root Attacks classes, best performance is related to high increasing class based on the decision tree.

Comparing the proposed method with the method of space of feature mapping with evolutionary calculations

It has been tried in this method to map the space of current features to the new space. In this method, it has been tried to map the input space into a new multi-dimensional decision space, making the most distinction. To obtain this aim, a Multi-purpose genetic method is used to extract appropriate features. The objective in extracting the features is to create the most distinction between different classes so that the computational cost becomes minimal. By using evolutionary calculations, the dimensions of the decision space is Optimized and mapped to new space. Finally, with the help of a simple classifier with several classes, Classifier is done. (Badran and Rockett, 2012)

In order to compare the proposed method And the two-step method, using multi-objective genetic programming, performance evaluation criterion including readout, precision rate and F criteria for standard class, Denial of Service Attacks , Probing , Remote to User Attacks and User to Root Attacks mapping of the Feature space was Calculated and shown in Table (4). Also, in Fig. 1 F evaluation criterion is shown as a chart.

Table 4: Comparing the results of the proposed method with feature space mapping method using evolutionary calculations.

F-Value	*Precision*	*Recall*	Method	Class
0.844561	0.733648	0.994983	[2](Rockett, 2012 & Badran) MMOGP	*Normal*
0.846367	0.736468	0.994818	Proposed Approach	
0.9844	0.999328	0.969911	MMOGP	*DOS*
0.985401	0.998887	0.972274	Proposed Approach	
0.730255	0.686378	0.780125	MMOGP	*Prob*
0.834636	0.849577	0820211	Proposed Approach	
0.105459	0.929158	0.055902	MMOGP	*R2L*
0.127149	0.808023	0.069003	Proposed Approach	
0.183099	0.464286	0.114035	MMOGP	*U2R*
0.346154	0.529412	0.257143	Proposed Approach	

[2] Multi-class pattern classification using single, multi-dimensional feature-space feature extraction evolved by multi-objective genetic programming and its application to network intrusion detection

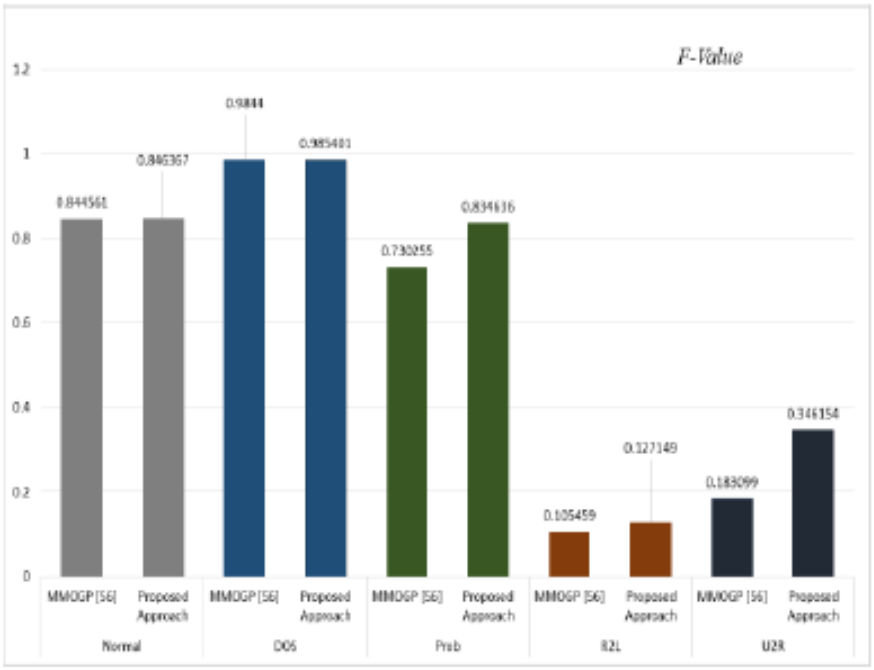

Figure 1: Comparison of F criterion of the results of proposed method with feature space mapping method by Evolutionary calculations

As shown in Table (4) and Figure (1), in all classes, the proposed method has improved F evaluation criterion which is the geometric combination of the two readout and precision criteria. This improvement is significant in both probing and User to Root Attacks classes. It has been improved 2 percent in Remote to User Attacks. There is not sensible improvement in standard behavior and Denial of Service Attacks because of high occurring samples and proper dispersion and proximity of the tutorial samples to the test samples.

The reason for a significant difference in the improvement rate of probing and User to Root Attacks in the proposed method to feature space mapping method by using evolutionary calculations can be mentioned as the following, because the tutorial data in dataset is completely unbalanced in probing and User to Root Attacks and standard behavior and Denial of Service Attacks, therefore methods like incremental algorithm based on The decision tree, functions better than feature selection based on evolutionary calculations which its main purpose is to function With unbalanced data.

CONCLUSION

The results have shown that this method had a better performance in all classes than the previously introduced methods. As the F criterion in standard lasses is equal to 0.8463, Denial of Service Attacks 0.9845, probing 0.8346, Remote to User Attacks 0.1271 and User to Root Attacks 0.3461. In the table (5) Comparison of F evaluation criterion in the proposed method and the displayed space mapping method.

Table 5: Comparison of F evaluation criteria of the proposed method and the two former methods based on five classes

Proposed Approach	MMOGP(Badran & Rockett, 2012)	Criteria	Class
0.846367	0.844561	*F-Value*	*Normal*
0.985401	0.9844	*F-Value*	*DOS*
0.834636	0.730255	*F-Value*	*Prob*
0.127149	0.105459	*F-Value*	*R2L*
0.346154	0.183099	*F-Value*	*U2R*

In the evaluation of the proposed combined method in two classes, the results also show that the proposed method could obtain detection precision of 0.917867 However, it was able to reduce the mistake alert rate by the amount of 0.005. Also, in the proposed method we obtained 0.95655 in F evaluation criterion which is the Geometric average of readout and precision.

Table 6: Comparison of the results of the proposed method with the former three methods based on the two ordinary and attack classes

Error Rate	*F-Value*	*False Alarm*	Method
0.0713	0.9536	0.005	MMOGP(Badran & Rockett, 2012)
0.067	0.9565	0.005	Proposed method

In Table 6 you can see an Overview of Evaluating the effectiveness of the proposed method Results Compared to the previous proposed two -class method. Of the most important factors in improving the results in the proposed method, some cases Such as feature selection, Sample selection, use of various and diverse classifiers to create mean and type of the final used classifier could be mentioned. The reasons for the weakness in detecting some classes include high imbalance of samples in the tutorial set, high distance and low topological similarity of tutorial and testing samples in some classes, existence of some sub- classes in the test set and lack of it in the tutorial set, high similarity of some classes to each other, and alignment of some weaknesses in classifiers and lack of proper coverage including sensitivity to noise data.

Reference

Badran, K., & Rockett, P. (2012). Multi-class pattern classification using single, multi-dimensional feature-space feature extraction evolved by multi-objective genetic programming and its application to network intrusion detection. Springer US Genetic Programming and Evolvable Machines Jornal, 33.

Catania, C., & Garino, C. (2012). Automatic network intrusion detection: Current techniques and open issues. Computers & Electrical Engineering, 1062-1072.

Chun, G., Yajian, Z., Ping, Y., Zhang, Z., Lio, G., & Yang, Y. (2014). A distance sum-based hybrid method for intrusion detection. Springer US Applied Intelligence Journal, 178-188.

Coolen, R., & Luiijf, H. (2002). Intrusion Detection: Generics and State-of- the-Art. Research and Technology Organization (RTO).

Depren, O., Topallar, M., Anarim, E., & Kema, M. (2007). An intelligent intrusion detection system (IDS) for anomaly and misuse detection in computer networks. Expert Systems with Applications, 713-722.

Hofman, A., Schmitz, C., & Sick, B. (2003). Rule extraction from neural networks for intrusion detection in computer networks. Systems, Man and Cybernetics. IEEE International Conference, 1259-1265.

Horng, S., Su, M., Chen, Y., & Kao, T. (2011). A novel intrusion detection system based on hierarchical clustering and support vector machines. Expert Systems with Applications, 306-313.

Kang, D., Fuller, D., & Honavar, V. (2005). Learning classifiers for Misuse and Anomaly Detection Using a Bag of system Calls representation. IEEE, Workshop on Information Assurance and Security.

Kaplantzis, S., & Mani, N. (2006). A study on Classification techniques for network intrusion detection. NC

Kayacik, H., Zincir-Heywood, A., & Heywood, M. (2003). On the capability of an SOM based intrusion detection system. Proceedings of the International Joint Conference on Neural Networks, 1808-1813.

Langely, P., Iba, W., & Thomas, A. (1992). An analysis of Bayesian classifier. Proceedings of the 10th National Conference on Artificial Intelligence 1992, Pages 223-228., 223-228.

Muda, Z., Yassin, W., Sulaiman, M., & Udzir, N. (2011). A K-Means and Naïve Bayes Learning Approch for Better Intrusion Detection. Information Technology Journal, 648-655.

Shelly Xiaonan, W., & Banzhaf, W. (2010). The use of computational intelligence in intrusion detection systems: A review. Applied Soft Computing, 1-35.

Xu, R., An, R., & Geng, X. (2011). Research intrusion detection based PSO-RBF classifier. Software Engineering and Service Science (ICSESS), IEEE 2nd International Conference, 104-107.