Fusion of Detectors on the Basis of Recirculation Neural Networks for Intrusion Detection

Pavel Kochurko
Brest State Technical University, Moskovskaja str. 267, 224017 Brest, Belarus, paulermo@tut.by

Abstract – The identification of attack class plays great role in intrusion detection. In this paper the method of recognition of a class of attack by means of the cumulative classifier with nonlinear recirculation neural networks as private detectors is described, strategy of detector selection – by a relative reconstruction error, relative cost of recognition error and mutual cost of recognition error are considered. Results of experiments are compared to results of similar researches.

Keywords – intrusion detection, classifier, recirculation neural networks, dynamic classifier selection

I. INTRODUCTION

Incessant distribution of application of information technologies to all spheres of human activity constantly puts new requirements to a level of security of information systems. Intrusion detection systems (IDS) already became a standard component of an infrastructure of network security. In spite of the fact that exist and constantly there are new methods of the analysis of network activity by means of various technologies of data mining [1], the basic technology of detection of attacks still is signature search. Its basic shortcoming – insufficient flexibility at detection of the modified attacks [2]. Considerably the best results at definition of the modified and new attacks are capable to show the systems using artificial neural networks [3-11]. Artificial neural networks (ANNs) have potential for the decision of a plenty of the problems covered by other modern approaches to intrusion detection. ANN have been declared alternatively to components of the statistical ana'ysis of systems of anomaly detection. Neural networks have been specially suggested to identify typical characteristics of users of system and statistically significant deviations from the established operating mode of the user [2].

In this paper the method of recognition of attack class on the basis of the analysis of the network traffic is described. Training and testing of ANNs was made on KDD'99 database which contains records describing TCP-connections including 41 parameter from processed DARPA 1998 Intrusion detection evaluation database [12]. The given data base includes normal connections, and also the attacks of 23 types belonging to four classes: DOS – «denial-of-service» - refusal in service, for example, a Syn-flood; U2R – not authorized access with root privileges on the given system, for example, various attacks of buffer overflow; R2L – not authorized access from the remote system, for example, password selection; Probe – analysis of the topology of a network, services accessible to attack, carrying out search of vulnerabilities on network hosts.

Paper is organized as follows. In section 2 variants of IDS architecture are described. In section 3 principles of application of the nonlinear recirculation neural networks (RNNs) for definition of an accessory of an entrance image to the given class are considered. Section 4 is devoted to application of fusion of classifiers on the basis of RNNs and a technique of the analysis and optimization of their teamwork. Conclusions are given in 5 section.

II. IDS STRUCTURAL ORGANIZATION

There are two basic technologies in intrusion detection: anomaly detection and misuse detection. Their basic difference consists that at use of the first the normal behavior of the subject is known and deviations from this behavior are searched while at use of the second attacks which are searched and distinguished among normal behavior. Both techniques eliminate each others defects, owing to what the best results of detection can be reached only applying them simultaneously (Figure 1), within the limits of different IDS subsystems [9] or with use of the combined detection methods [10].

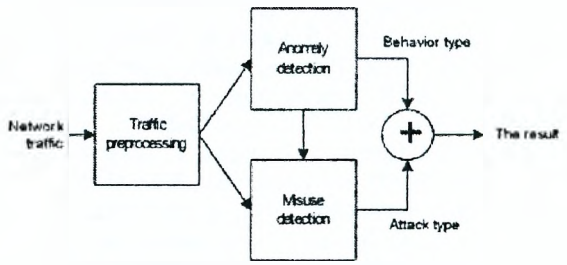

Fig. 1 – Simplified IDS structure with the anomaly detector and the block of recognition of attack

Acting on an input of system the network traffic passes preprocessing then data about network connections act on an input of the detector of anomalies and the block of recognition of attack. Thus on quality of work of the first depends – whether it will be found out, in fact if the detector of anomalies characterizes connection as normal, the result of recognition means is not so important. The method of its training is applied to improvement of finding out ability of the anomaly detector on the combined data set – normal connections and attacks [10] that leads to a combination in it of both technologies.

It is proved [13, 14], what the best results at classification (even a question – «attack or not?» is definition of an accessory to a class of attacks or a class of normal connections; not speaking already about definition of a class of attack) give classifiers independent from each other. The basic problem in system engineering from several independent detectors or classifiers becomes a question of a choice of the most plausible value among the results which are given out by different classifiers (a

dynamic classifier selection - DCS). In case of application of "too independent" detectors there is a danger, that construction of the general estimation will be complicated because of incommensurable or incomparable outputs of detectors. So, in case of application of RNNs as the anomaly detector and multilayered perceptron (MLP) as the misuse detector [9], it is possible to operate only with answers of detectors – attack or not – and any others more or less comparable characteristics (reconstruction error on the anomaly detector and values of MLP outputs are not comparable).

There are much more abilities for construction of a cumulative estimation of the general classifier at use of independent detectors of the identical nature. In this case outputs of each separate detector are comparable among themselves, also various DCS methods can be applied: an average estimation, the maximal vote, a "a posteriori" method, etc. [13], or as described in section 4.

III. RNNs BASED DETECTORS

A. The anomaly detector

Recirculation neural networks (Figure 2) differ from others ANNs that on the input information in the same kind is reconstructed on an output. They are applied to compression and restoration of the information (direct and return distribution of the information in the networks «with a narrow throat») [15], for definition of outliers on a background of the general file of entrance data [16].

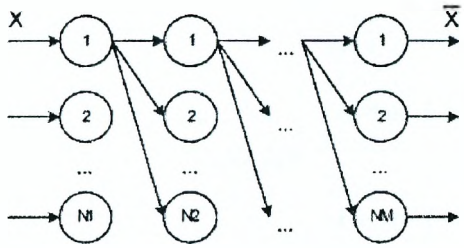

Fig. 2 – *M* layers RNN structure
***N_i*– quantity of neural elements in *i-th* layer, *NM=N1* – quantities of neural elements in entrance and target layers are equal**

Nonlinear RNNs have shown good results as the detector of anomalies [9, 10]: training RNN is made on normal connections so that input vectors on an output were reconstructed in themselves, thus the connection is more similar on normal, the less reconstruction error is:

$$E^k = \sum_j (\bar{X}_j^k - X_j^k)^2, \qquad (1)$$

where X_j^k – j-th element of k-th input vector, $\bar{X}_j^k$ – j-th element k-th output vector. Whether $E^k > T$, where T– certain threshold for given RNN connection admits anomaly, or attack, differently – normal connection. Thus there is a problem of a threshold T value determination, providing the most qualitative detection of abnormal connections. It is possible to get threshold value minimizing the sum of false positive (FP) and false negative (FN) errors, basing on cost characteristics of the given errors – FN error seems to be more expensive, than FP error, and its cost should be higher [10].

B. Private classifiers

The described technique of definition of an input vector accessory to one of two classes – "normal" or "attacks", that is "not-normal" – it is possible to use in opposite way. If at training the detector of anomalies we used normal vectors which were restored in itself, and the conclusion about their accessory to a class "normal" was made, training the detector on vectors-attacks which should be restored in itself, it is possible to do a conclusion about their accessory to a class of "attack". Thus, if during functioning of this detector the reconstruction error (1) exceeds the certain threshold, given connection it is possible to carry to a class "not-attacks", that is normal connections. As training is conducted on vectors-attacks the given approach realizes technology of misuse detection, and its use together with previous technique is righteous.

Thus, one RNN can be applied to definition of an accessory of input vector to one of two classes – to on what it was trained (class A), or to the second (class $\bar{A}$), to which correspond outliers:

$$\begin{cases} X^k \in A, \ \text{if} \ \ E^k \le T, \\ X^k \in \bar{A}, \ \text{if} \ \ E^k > T. \end{cases} \qquad (2)$$

Worth to note that is possible to train RNN in the special way [10] on connections of both classes so that to raise quality of detection on conditions (2).

As already it was mentioned above, database KDD includes normal connections and also attacks of four classes which considerably differ from each other. Therefore it is advisable to train detectors for each of five classes separately, not uniting all classes of attacks in a single whole.

Here again there is a problem of a choice of a threshold T for each concrete detector. If for the anomaly detector it was possible to speak at once, that cost of FN error is higher, than cost of FP error, in case of the detector for a class of attacks R2L it is herd to tell what will be worse – FP error (that is to name "R2L" connection to this class not concerning – attack of other class or normal connection) or FN detection of the given attack (on the contrary).

Many researchers [17] use a cost matrix for definition of cost of errors F (Table 1). Average values of FP and FN errors for each class can be calculated as follows:

$$F_i^{FP} = \frac{\sum_{j,\ j \ne i} F_{ji}}{N-1}, \quad F_i^{FN} = \frac{\sum_{j,\ i \ne j} F_{ij}}{N-1}, \qquad (3)$$

Table 1. The cost matrix F of incorrect classification of attacks

Real class	Prospective class				
	normal	dos	probe	r2l	u2r
1 normal	0	2	1	2	2
2 dos	2	0	1	2	2
3 probe	1	2	0	2	2
4 r2l	4	2	2	0	2
5 u2r	3	2	2	2	0

where N – quantity of classes *(N=5)*. Proceeding from the given matrix it is possible to draw a conclusion, that FP error for the detector of a class "normal" on the average has cost 2,5 (the sum of elements of a column "normal" divided by 4), and average FN error will cost 1,75 (the sum of elements of a line "normal" divided by 4). As FP error of the detector of a class "normal" is as a matter of attack undetection, that is FN error of all system, and FN error of the detector of a class "normal" – false detection (FP error) of all system, the given parity repeats told above, that FN errors of system are more expensive than FP.

It is similarly possible to calculate average costs of errors F_i^{FP} and F_i^{FN} for $\forall i \in [1..5]$ - that is for detectors of all classes (Table 2).

On the basis of the given costs it is possible to choose value of a threshold which minimizes a total average error on training or validation data base.

C. Experimental results

For an estimation of efficiency of the offered approach a number of experiments is lead. Private detectors for each class are trained, and all over again the training set got out of all base KDD, then from connections on concrete services – HTTP, FTP_DATA, TELNET. Nonlinear RNNs were used with one hidden layer with function of activation a hyperbolic tangent and logical sigmoid function of activation in a target layer. Quantity of neural elements in input and target layers according to quantity of parameters of input data – 41, in the hidden layer – 50.

After each detector was trained the testing on training samples was conducted with the purpose of a finding of value of threshold T at which average cost of an error is minimal. In the further the testing of trained detectors was made on test samples with threshold values received before (Table 3).

IV. FUSION OF PRIVATE CLASSIFIERS

A. Joint functioning

As it was told above the best classification results can be achieved using several independent classifiers of the identical nature, because construction of the general estimation from private can be made by greater number of methods. We shall unite the private detectors trained in the previous section in one general (Figure 3).

The basic problem in construction of such classifier becomes definition of a cumulative estimation proceeding from estimations of private detectors. In works of various researchers (for example [13]) the set of methods, such as a finding of average value for each class on the basis of indications of all classifiers, the sum of votes for each

Table 2. Average costs of errors of detectors of each class

Class	Cost	
	F_i^{FP}	F_i^{FN}
1 normal	2,5	1,75
2 dos	2	1,75
3 probe	1,5	1,75
4 r2l	2	2,5
5 u2r	2	2,25

Table 3. Results of testing of detectors

Service	Threshold	FP, %	FN, %	Average cost
ALL				
normal	0,00070	12,56	6,68	0,1844
dos	0,00214	4,33	1,09	0,0542
probe	0,00120	7,79	14,21	0,1675
r2l	0,00116	2,87	5,38	0,0947
u2r	0,00112	7,07	5,54	0,1323
HTTP				
normal	0,00620	2,4	0,17	0,0214
dos	0,00290	1,5	0	0,0098
probe	0,00114	0	0	0
r2l	0,00110	0	0	0
FTP DATA				
normal	0,00123	6,8	2,72	0,0841
dos	0,00340	0	0	0
probe	0,00132	0	0	0
r2l	0,00114	5,17	0,25	0,0463
u2r	0,00126	0	0,07	0,0009
TELNET				
normal	0,00036	44,4	1,31	0,2394
dos	0,00650	0	0	0
probe	0,00162	0	0	0
r2l	0,00136	3,33	0	0,0294
u2r	0,00076	5,91	2	0,0907

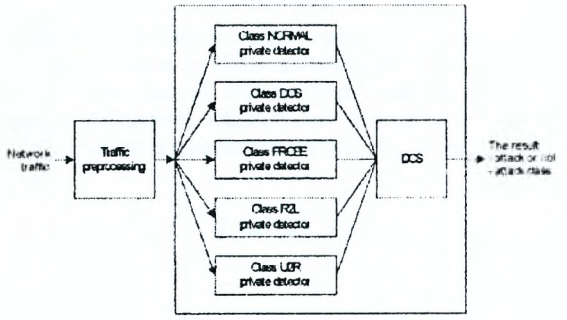

Fig. 3 – Fusion of independent private classifiers in one general

class, methods of an estimation «a priori» and «a posteriori» is considered. These methods mean that each classifier states a private estimation concerning an opportunity of an accessory of input image to at once several classes, and these classes are identical to all classifiers. However in our case classes, about an

accessory to which each classifier judges, first, are various, secondly, are crossed. Therefore all the methods listed above are not applicable.

B. Dynamic classifier selection

The general classifier consists from $N=5$ private detectors, each of which has a threshold T_i. Values of thresholds got out proceeding from minimization of average cost of errors. To make estimation values comparable it is enough to scale reconstruction error on a threshold. Then (2) will be:

$$\begin{cases} X^k \in A_i, \ \ if \ \ \delta_i^k \le 1, \\ X^k \in \overline{A}_i, \ \ if \ \ \delta_i^k > 1, \end{cases} \tag{4}$$

where $\delta_i^k = \frac{E_i^k}{T_i}$ - a relative reconstruction error. Thus, than less δ_i^k, the probability of accessory of an input image X^k to a class A_i is higher. Therefore it is possible to allocate the first method of determination of a cumulative estimation – by *the minimal relative reconstruction error*:

$$\begin{cases} X^k \in A_m, \\ \delta_m^k = \min\limits_i \delta_i^k. \end{cases} \tag{5}$$

As the purpose of improvement of efficiency of classification is the minimization of erroneous classification expressed in minimization of average cost of classification, in construction of a cumulative estimation it is possible to act the same as at the choice of a threshold in private detectors – to consider cost of erroneous classification. If δ_i^k - a characteristic of probability of error of classification on *i-th* detector the estimation of possible average cost of error on each of detectors will be equal:

$$\Omega_i^k = \frac{\sum\limits_{j, j \ne i} \delta_i^k F_{ji}}{N-1}. \tag{6}$$

The estimation (6) shows, what ability of loss in cost if we shall name a vector belonging to *j-th* class by a vector of *i-th* class, i. e. *i-th* classifier instead of *j-th* will be chosen. On the basis of the given estimation we shall allocate the second method of a cumulative estimation determination – on *the minimal possible cost of false classification*:

$$\begin{cases} X^k \in A_m, \\ \Omega_m^k = \min\limits_i \Omega_i^k. \end{cases} \tag{7}$$

Besides it is possible to consider mutual influence of possible errors – to add up an estimation Ω_i^k and an estimation of a prize in cost if *i-th* classifier instead of wrong *j-th* will be chosen:

$$\Psi_i^k = -\frac{\sum\limits_{j, j \ne i} (\delta_j^k - \delta_i^k) F_{ij}}{N-1}. \tag{8}$$

Then on the basis of estimations (6) and (8) it is possible to allocate the third rule of winner detector selection – on *the minimal possible mutual cost of false classification*:

$$\begin{cases} X^k \in A_m, \\ \Omega_m^k + \Psi_m^k = \min\limits_i (\Omega_i^k + \Psi_i^k). \end{cases} \tag{9}$$

C. Experimental results

Efficiency of the general classifier functioning we shall check up experimentally in such way as private detectors from section 3. Results are presented in Table 4.

Table 4. Results of detection and recognition of attacks by the cumulative classifier

DCS	FP, %	FN, %	Quality of recognition				Av. cost
			dos, %	probe %	r2l, %	u2r, %	
ALL							
(5)	10,8	2,3	98,2	96,6	91,9	100	0,061
(7)	30,8	0,9	97,8	99,3	92,5	100	0,076
(9)	18,8	0,7	98,3	98,0	93,1	98,2	0,074
HTTP							
(5)	0	0,1	99,8	100	100	–	0,001
(7)	0	0,1	99,8	100	0	–	0,287
(9)	0	0,1	99,8	100	100	–	0,001
FTP DATA							
(5)	0,7	1,1	100	100	96,7	100	0,043
(7)	0,7	1,1	100	100	96,7	100	0,043
(9)	27,3	0,4	100	77,6	98,7	100	0,173
TELNET							
(5)	0	5,3	98,8	100	97,3	85,5	0,150
(7)	0	5,0	97,8	100	98,0	85,5	0,145
(9)	15,0	1,6	98,5	100	98,0	96,9	0,068

Apparently from results, the unequivocal answer to a question – what method is better – is not present. The method of a choice of a final class with use of mutual cost (9) can minimize a error, but with substantial growth of quantity of false detection (FP), methods (5) and (7) give basically comparable results, on some service one is better, on some – another.

V. CONCLUSIONS

Let's compare results which have shown experiments with use of the described technique and the results received within the other researches (Tables 5 and 6).

Comparing values in tables 4-6, it is possible to note, that quality of detection of attacks by the described technique does not concede (at application of one classifier for all

services) and considerably surpasses (at application of separate classifiers for each service) analogues. The level of recognition of classes of attacks has considerably improved results shown earlier (RNN+MLP), especially for attacks of classes r2l and u2r.

Table 5. Results of detection by means of various technologies [8]

Technology	FN, %	FP, %
Data mining [18]	10-30	2
Clusterisation [19]	7	10
K-NN [19]	9	8
SVM [19]	2	10

Table 6. Results of recognition of classes of attacks in some researches

	dos, %	probe, %	r2l, %	u2r, %
KDD-99 Winner [20]	97,12	83,32	13,16	8,40
SOM [8]	96,70	79,70	18,40	30,00
RNN+MLP [11]	99,98	98,78	45,20	3,84

The shortcomings of the given technique which it is necessary to work on in the further: strong dependence of quality of detection on threshold values of private detectors. Values of thresholds are determined proceeding from cost parities which base on an expert estimation, therefore construction of techniques of determination of the best values of thresholds only will improve quality and stability of work of system.

Thus, it is possible to draw a conclusion, that the method of the cumulative classifier on the basis of nonlinear recirculation neural networks as private detectors can be applied with success to the solving of problems of recognition of network attacks and other problems of recognition of images.

ACKNOWLEDGMENT

This research is supported by the grant of Belarus National Academy of Sciences and the grant of Belarus Ministry of Education.

REFERENCES

[1] Brugger S. T. Data Mining Methods for Network Intrusion Detection. http://www.bruggerink.com/~zow/Projects.html

[2] A. V. Lukatsky. Intrusion detection. – Saint-Petersburg: BHV-Peterburg, 2003.

[3] J. Cannady. Applying Neural Networks to Misuse Detection. In *Proceedings of the 21st National Information Systems Security Conference.*

[4] J. M. Bonifacio et al. Neural Networks applied in intrusion detection systems, In *Proc. of the IEEE World congress on Comp. Intell.* (WCCI'98), 1998.

[5] C. Jirapummin and N. Wattanapongsakorn. Visual Intrusion Detection using Self-Organizing Maps. In *Proc. of Electrical and Electronic Conference (EECON-24)*, Thailand, Vol. 2, pp. 1343-1349, 2001.

[6] D. Joo, T. Hong and I. Han. The neural network models for IDS based on the asymmetric costs of false negative errors and false positive errors. *Expert Systems with Applications, 25 (2003),* pp. 69-75

[7] C. Zhang, J. Juang, M. Kamel. Intrusion detection using hierarchical neural networks. *Pattern Recognition Letters (2004).*

[8] H. G. Kayacik. Hierarchical self organizing map based IDS on KDD benchmark. M. Sc. work, Dalhousie university, Halifax, Nova Scotia, 2003.

[9] V. Golovko, P. Kochurko. Some Aspects of Neural Network: Approach for Intrusion Detection. In Kowalik J., Gorski J., Sachenko A. editors, *Cyberspace Security and Defense: Research Issues* Springer, 2005, VIII – pp. 367-382

[10] P. Kochurko, V. Golovko. Neural Network Approach to Anomaly Detection Improvement. In Proc. of 8th International Conference on Pattern Recognition and Information Processing (PRIP'05), May, 18-20, Minsk, Belarus, 2005 – pp. 416-419.

[11] V. Golovko, P. Kochurko. Intrusion recognition using neural networks. *International Scientific Journal of Computing*, vol.4, issue 3, 2005, p.37-42

[12] KDD Cup'99 Competition, 1999, http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

[13] Giacinto G., Roli F., Fumera G. Selection of image classifier. *Electron*, 26(5), 2000, pp. 420-422.

[14] Xu L., Krzyzak A., Suen C. Y. Methods for combining multiple classifiers and their applications to handwriting recognition. *IEEE Trans. Syst. Man Cybernitics*, 22, 1992, pp. 418-435

[15] V. Golovko, O. Ignatiuk, Yu. Savitsky, T. Laopoulos, A. Sachenko, L. Grandinetti. Unsupervised learning for dimensionality reduction. *Proc. of Second Int. ICSC Symposium on Engineering of Intelligent Systems EIS'2000*, University of Paisley, Scotland, June 2000. Canada / Switzerland: ICSS Academic Press, pp. 140 – 144, 2000

[16] S. Hawkins et al. Outlier Detection Using Replicator Neural Networks. In *Proc. of the 4th International Conference on Data Warehousing and Knowledge Discovery (DaWaK02) Lecture Notes in computer Science*, Vol. 2454, Springer, Pages 170-180, ISBN 3-540-44123-9, 2002

[17] Giacinto G. et al. Fusion of multiple classifiers for intrusion detection in computer networks. *Pattern Recognition Letters,* 24, 2003, pp. 1795-1803

[18] Lee W., Stolfo S. A Framework for Constructing Features and Models for Intrusion Detection Systems. *Information and System Security,* 3(4), 2000, pp. 227-261

[19] Eskin E. et al. A Geometric Framework for Unsupervised Anomaly Detection: Detecting intrusiond in unlabeled data. In D. Barbara and S. Jajodia editors, *Applications of Data Mining in Computer Security.* Kluwer, 2002.

[20] Pfahringer B. Winnings the KDD99 Classification Cup: Bagged Boosting. *SIGKDD Explorations,* 1(2), 2000, pp. 65-66