2005 IEEE International
Symposium on Signal Processing
and Information Technology

Detecting Denial of Service Attacks Using Emergent Self-Organizing Maps

Aikaterini Mitrokotsa, Christos Douligeris
Department of Informatics, University of Piraeus,
80 Karaoli and Dimitriou Str. Piraeus 18534, Greece
{mitrokat, cdoulig}@unipi.gr

Abstract - *Denial of Service attacks constitute one of the greatest problem in network security. Monitoring traffic is one of the main techniques used in order to find out the existence of possible outliers in the traffic patterns. In this paper, we propose an approach that detects Denial of Service attacks using Emergent Self-Organizing Maps. The approach is based on classifying "normal" traffic against "abnormal" traffic in the sense of Denial of Service attacks. The approach permits the automatic classification of events that are contained in logs and visualization of network traffic. Extensive simulations show the effectiveness of this approach compared to previously proposed approaches regarding false alarms and detection probabilities.*

Keywords *– Emergent Self Organizing maps, Denial of service, Intrusion detection, Neural Networks*

I. INTRODUCTION

The explosive growth of computer networks and particularly of the Internet has created many stability and security problems. One of the greatest threats that network security faces nowadays is Denial of Service attacks. The need for a defense against Denial of Service attacks is becoming an important challenge and a difficult task at the same time.

Intrusion detection systems have been developed in order to defend computer networks against the continuous evolution of various types of threats, including Denial of Service attacks. Intrusion detection techniques can in general be divided into anomaly detection and misuse detection.

Misuse detection uses a priori knowledge on intrusions and tries to detect attacks based on specific patterns or signatures of known attacks. Although misuse detection systems are very accurate in detecting known attacks, their basic drawback is that network attacks are under a continuous evolution and this leads to the need for an up-to date knowledge base of all attacks. Anomaly detection has the advantage of being able to detect unknown attacks. Anomaly detection techniques define what is normal and attempt to track deviations from the normal behaviour that are considered to be intrusions.

In this paper we describe a network-based anomaly detection method in order to detect Denial of Service attacks. In order to relieve network administrators from the extremely difficult and time consuming task of scanning network traffic and to avoid large processing overheads, we propose an anomaly detection approach that is based on a class of neural networks known as Kohonen's Self-Organizing Maps (KSOMs).

Neural networks have the great advantage of tolerance in imprecise data. Emergent SOMs are based on simple Kohonen's SOMs but present some advantages over them that can be exploited in order to achieve better results in the detection of intrusions. Combining machine learning and information visualization techniques we may be able to have a clearer view of how secure is our network against DoS attacks.

The proposed approach can be used either for analyzing real-time data or past network traffic. For the evaluation of our approach we used the KDD-99 [1] dataset and not ad hoc generated data in order to ensure the reliability of our results.

Following this introduction, the paper is organized as follows. Section 2 describes the SOM algorithm. Section 3 presents related work of intrusion detection approaches that have been proposed and rely on SOMs, while section 4 describes Emergent SOMs and their differences from simple Kohonen's SOMs. In section 5 we present our approach that is based on Emergent SOMs for the detection of DoS attacks. We also present the evaluation of our approach and the experimental results. Section 6 concludes the paper and discusses some future work.

II. SELF-ORGANIZING MAPS

Kohonen's Self-Organizing maps [2] have their base in biology. They belong in the category of unsupervised or competitive learning network and produce a 2D topological map, which illustrates the input data according to their similarity.

They are unsupervised, since there are no target vectors as in the case of back propagation vector networks. The Self Organizing Map is trained using only the characteristics of the trained data. They are competitive, because there is only one winning neuron in the output layer. That is why the Self Organizing Map is also referred as a winner-take-all unsupervised learning neural network.

0-7803-9314-7/05/$20.00©2005 IEEE

The trained KSOMs create clusters of data, where similar vectors of features are located in a specific region in the output space. This is very useful for discovering clusters and relationships in data. The generated mapping is topology preserving. The KSOMs are extensively used for data analysis and visualization.

The learning procedure is composed of the following steps:

a. The random weights w_{ij} (also known as codebook vectors of the neurons) are initialized with small random values.
b. Use an input pattern x
c. Calculate the Euclidean distance (eq. 1 [2]) between input data sample x, and each neuron weight w_{ij}. The winner (Best Matching Unit) is chosen as o(x):

$$o(x) = \arg\min_{j} \left\| x - w_{ij} \right\|, \text{ j=1,2,}\ldots l \quad (1)$$

d. In order to achieve the topological mapping, all the weights in the neighborhood are adjusted, depending on their distance from the winning neuron according to the following equation [2]:

$$\forall j : w_{ij}(t) = w_{ij}(t-1) + a(t)\eta(t') \cdot \left(x_i(t) - w_{ij}(t-1) \right) \ (2)$$

where α is the learning rate and η the neighborhood function and *t'*, the time that was spent in the current context. The neighborhood function η decreases as *t'* increases.

e. Repeat steps b, c, d until convergence.

Self-organizing maps are trained until convergence, which depends on how complex is the input dataset. SOMs permit a better overview of the data that represent the network traffic. Furthermore, it is a technique that gives the advantage of finding hidden relations of data and segments them visually, something that leads to a much easier visualization of data.

SOMs require minimal expert knowledge and achieve high speed and high conversion rates in comparison with other learning techniques without extensive off-line training. Kohonen' s SOMs present many limitations that make a single Kohonen map unable to reliably characterize disparate information like network traffic. Emergent SOMs can help us produce more reliable results.

III. SELF-ORGANIZING MAPS AND INTRUSION DETECTION

KSOMs have been used extensively in the area of Intrusion Detection. Girardin et al. [3] propose an experimental system, which is based on spring layouts and unsupervised neural networks that can be used for the classification of event logs, usage assessment, real-time trend analyses, anomaly detection and break-in attempts detection. Nguyen [4] describes a software package called iSOM, a module of the intrusion detection system INBOUNDS. This software package uses SOM in order to detect anomalies in computer network traffic.

Hoglund et al. [5] propose a prototype UNIX anomaly detection system, which is host based and monitors computer network host users, based on SOMs to test if user behavior is anomalous. Lichodzijewski et al. [6] apply hierarchical SOMs in order to achieve host based intrusion detection. The hierarchical SOM architecture that is used consists of two levels. The first level consists of three maps. Each map visualizes one attribute of data vectors and time. The map of the second level combines the results form each first level map in order to provide an integrated view of the network state.

K. Labib et al. [7] propose an anomaly detection system called NSOM, which attempts to classify real-time Ethernet data using Kohonen SOMs. The KSOM is trained with normal data and then the system is tested with real-time Ethernet data. An attack is detected if the winning neuron is not one of the noted neurons. Rhodes et al. [8] propose an approach for intrusion detection using multiple Self-Organizing maps. Their approach is not based on a single SOM to detect intrusions but they construct a monitor stack architecture with multiple SOMs, each becoming a specialist on recognizing abnormal activity in a specific protocol.

Jirapummin et al. [9] propose an approach that is based on SOM and Resilient Propagation Neural Network (RPROP) in order to achieve intrusion detection combined with visualization (by SOM) and classification (by RPROP) on normal traffic and intrusions. Gonzalez and Dasgupta [10] use KSOMs in order to compare the results of a Neuro-Immune approach for anomaly detection they have proposed with an unsupervised technique (SOM). This comparison does not lead to a clear winner.

Kayacik et al. [11] propose an approach for Intrusion detection that is based on a hierarchy of KSOMs. They try to define how far an intrusion detection approach using a sequence of hierarchical SOMs using only 6 features from the 41 features of KDD dataset.

All the above presented intrusion detection approaches are based on simple KSOMs and can be further improved if Emergent Self-Organizing Maps are used.

IV. EMERGENT SELF-ORGANIZING MAPS

Although as we have described in the previous section KSOMs present many advantages and have been extensively used in the research area of intrusion detection a special category of Self Organizing Maps called Emergent Self-Organizing Maps seem to provide much more advantages against Kohonen's SOM that can be exploited in order to achieve better results in the Intrusion Detection process.

According to [12] “Emergence is the ability of a system to produce a phenomenon on a new, higher level”. In order to achieve emergence the existence and cooperation of a great number of elementary processes is necessary. Emergence may be presented not only in natural but also in technical systems. Emergent SOMs has been applied in a

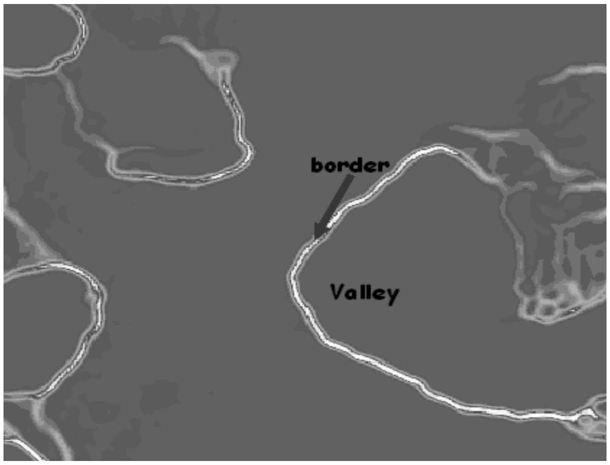

Fig. 1. Emergent SOM U-Matrix

broad range of areas including medical diagnosis and environmental science.

One of the basic disadvantages of SOM maps is that their abilities are limited to a few neurons. The small number of neurons (some tens) of Kohonen's SOMs does not permit them to show emergence. More specifically in a Kohonen's SOM to each neuron correspond the best matches of a great number of input data. So in a way each neuron represents a cluster. The limitation of few neurons limits the topology preservation of Kohonen's SOMs to small maps. The clustering that is based on a Kohonen's SOM has many similar points to the clustering algorithm k-Means. More specifically a SOM with a few neurons is almost the same with the clustering algorithm k-Means [13] when k is the number of neurons of the SOM.

On the other hand, Emergent Self-Organizing Maps may expand from some thousands to tens of thousands of neurons. More specifically the number of neurons in some cases may be greater than the number of input data. As a result only a small number of input data may correspond to each neuron.

The large number of neurons in ESOM is necessary in order to achieve emergence. The cooperation of such a big number of neurons leads to structures of a higher level. This cooperation permits to observe systems in a higher level observing the overall structures, disregarding the elementary ones and allowing to consider structures that otherwise would be invisible. The clustering procedure in emergent SOMs is performed by observing the whole Emergent Self-Organizing Map and not by focusing on its neurons.

There are many methods that can be used in order to visualize the structures generated by ESOMs including distance-based (U-Matrix), density-based (P-Matrix), distance and density based (U*-Matrix) and topology visualizations. We have used the distance based (U-Matrix) method. According to this method [12] the sum (height) of distances between the neuron-weights and its neighbours is normalized by the largest height. The result of the sum of distances is represented as elevation of each neuron, this way the input data set is displayed and depicted as a 3D landscape. The height will have a large value in areas of the map where few datapoints belong and small in areas that represent clusters. Thus hills and valleys correspondingly will be created (fig. 1). The height (*$uh(n_i)$*) of each neuron (*n_i*) is given from the following equation (eq.3) [13]:

$$uh(n_i) = \sum_{n_j \in U_i} d(n_i, n_j) \quad (3),$$

where U_i represent the neighbor neurons of n_i.

V. INTRUSION DETECTION WITH ESOM

We trained Emergent SOMs with logs of network traffic and exploited the main advantage of ESOMs the large number of neurons. In order to visualize these structures the U-Matrix method is used. This method permits us to achieve a good visualization of the network traffic and observe the existence of possible intrusions.

We used ESOM U-matrices [13] in order to perform Denial of Service detection. In our case, each log of network traffic is represented by a vector with some fixed attributes. Each vector has a unique spatial position in the U-Matrix and the distance between two points is the dissimilarity of two network traffic logs. The U-Matrix of the trained dataset is divided into valleys that represent clusters of normal or attack data and hills that represent borders between clusters.

Depending on the position of the best match of an input data point that characterizes a connection this point may belong to a valley (cluster (normal or attack behaviour)) or this data point may not be classified if its best match belongs to a hill (boundary). The map that will be created after the training of the Emergent SOM, will represent the network traffic. Thus an input data point may be classified depending on the position of its best match.

In order to avoid having a great of influence of the attributes of some input vectors it is necessary to normalize the input data. Many methods are used for the data normalization. We have normalized the data with mean zero and variance one, a technique that produces very good results in most cases as reported in the literature.

A. The KDD Dataset

The KDD-99 (Knowledge Discovery in Databases) [1] dataset is a standard set of data that can be used in order to evaluate proposed approaches in the area of Intrusion Detection. The attacks included in the dataset belong to one of the following categories: DoS, R2L (Remote to Local), U2R (User to Root) and probing. In this dataset, every connection is characterized by 41 features.

We have focused on the detection of DoS attacks [14]. DoS attacks constitute one of the major threats and among the hardest security problems in computer networks. The main aim of a DoS attack is the disruption of services by attempting to limit access to a machine or service. Depending on the attackers' strategy, the target resources may be the file system space, the network bandwidth, or the network connections, resulting in a network incapable of providing normal service. For the evaluation we have used the Databionics ESOM tool ([15], [16]).

B. Selection of Features

A very important decision is the selection of feature vectors that would be used in the ESOM classification. The features of network traffic should be in a suitable form in order to be easily processed by ESOM and representative of network activity in order to be able to distinguish normal and abnormal activity. It is important that the selected features of the vectors increase the contrast between normal and abnormal activity concerning DoS attacks.

Mukkamala et al. [17] identified the most significant features from the KDD-99 dataset using two ranking methods for the SVMs (Support Vector Machines) and ANNs (Artificial Neural Networks). According to [20] the most important features for the detection of Denial of Service attacks from the 41 features of KDD data are the following:

- Duration
- Source bytes
- Destination bytes
- Count
- Same service rate
- Connections with SYN errors
- Connections-Same service-SYN errors
- Destination-Host-SYN error rate
- Destination-Host-Same-Service error rate

We used these features in our approach considering the fact that they are the most important for the detection of DoS attacks.

C. Experimental Results

The presented evaluation proves that we can achieve an accurate differentiation between normal and abnormal behaviors concerning DoS attacks. In order to perform clustering with ESOM U-Matrices we followed the proceeding procedure. The best matches of the trained dataset and thus the corresponding dataset are manually grouped into clusters representing normal and attack behaviour. Thus we identify the regions of the map that represent a cluster that can be used for the classification on new datasets.

We performed various evaluation experiments. Table I presents the datasets that were used in order to train and test our approach. All the datasets are parts of the available 10% KDD dataset of various sizes that include normal data and DoS attacks (smurf, Neptune, pod, back, teardrop). We have to note here that because the training procedure training procedure presents a high computational overhead especially if the number of input data set is over 10.000, the evaluation experiments where limited to datasets whose size ranges from about 20.000 to 60.000 records. Of course this disadvantage can be balanced if we consider that the training is performed only once and the accuracy of results for the test data is very high. The ESOM of a trained dataset can be depicted in figure 2. As it can be clearly seen the training data set has been divided in two classes that are very well distinguished, normal data class (dark color) and DoS data class (light color). In figure 3 is depicted the testing dataset for the corresponding training dataset.

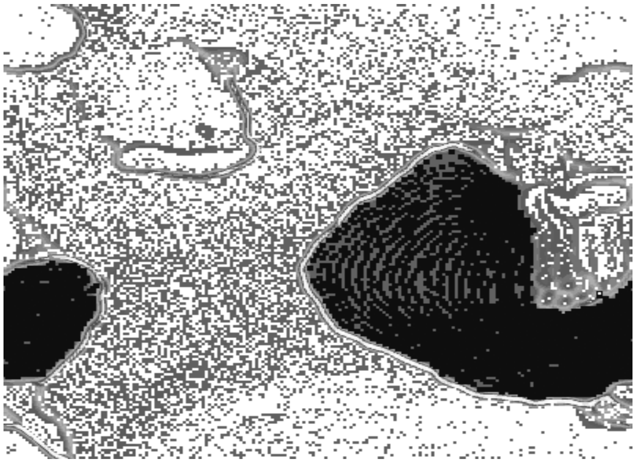

Fig. 2. Emergent SOM U-Matrix of train dataset

In order to evaluate the efficiency of the proposed approach we use two measures: the detection rate and the false alarm rate that has been defined in [8], as follows:

$$\text{Detection rate} = \frac{TP}{TP+FN}, \text{ False alarm rate}= \frac{FP}{TN+FP},$$

where TP is the number of true positives (attack logs classified as attacks), TN the number of true negatives (normal logs classified as normal), FP the number of false positives (normal logs classified as attacks) and FN the number of false negatives (attack logs classified as normal). The most effective approach should reduce as much as possible the *False alarm rate* and at the same time increase the *Detection rate.*

Table II presents the results of the evaluation of the various experiments we have performed. In this table the detection rate and the false alarm rate for each experiment are presented as well as the training parameters that were used for the ESOM training. According to [15] in order to avoid topology errors the identical ESOM architecture should have at least 4000 neurons and the ratio rows and columns should be different from unity. Thus experiments

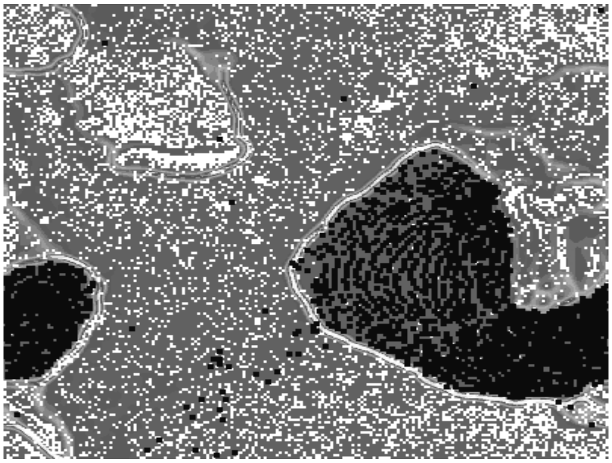

Fig. 3. Emergent SOM U-Matrix of test dataset

have been performed either with 160x180 or 180x 200 neurons and the training epochs range between 50 and 60, the Gaussian function was used as kernel neighborhood function and weight initialization method and the Euclidean as distance function. The initial and final learning rate were 0.5 and 0.1 respectively and the initial value for radius were 79 (for 160x180 neurons) and 89 (for 180x200) and 1 the final radius. Moreover in order to avoid topology errors caused by border effects we have used boundless toroid grids.

The most recent work in this area using SOM [13] that outperforms all previous proposals is able to provide detection rates for all attacks included in KDD dataset, in the range of 89% to 99.7% depending on the KDD-99 test partition employed and false positive rate (false alarm) ranging between 4.6% and 1.7% respectively. In our approach because of focusing in Denial of Service attacks we have used 9 features instead of 6 that have been used in [13]. The evaluation experiments of our approach provide detection rate for DoS attacks that ranges between 98,3% and 99,81% and the false alarm that ranges between 2.9% to 0.1%. For the evaluation of our approach we have used the KDD dataset that is widely used in order to compare intrusion detection approaches as a benchmarking dataset. Although the subsets we have used are randomly selected and different from subsets used in previous approaches the results are extremely promising. From the above results we see that the proposed method achieves extremely good results and has a lower deviation from the other proposed methods.

Moreover, in order to make sure that our approach will always provide efficient and accurate results we should update our trained ESOM U-matrix according to new DoS attack patterns.

VI. CONCLUSIONS AND FUTURE WORK

By exploiting the visualization of network traffic our approach detects Denial of Service attacks by classifying malicious and normal actions. The proposed approach is extremely powerful in producing efficient results. Its main advantage lies in the fact that Emergent SOMs extend the abilities of simple KSOMs by developing high-level structures that could be invisible with simple KSOMs where only a few neurons can be used.

The main disadvantage is the high computational overhead caused during training datasets that have greater size than 10.000 records. But certainly ESOM's computational overhead does not prohibit its use. Moreover, during the classification procedure the classes of the trained data have to be defined manually through the observation of the map that may introduce a process error. Although we should note that for our experiments we employed small datasets of 10% KDD dataset the results are extremely promising.

Even though, in this paper we focused on DoS attacks we plan to examine the performance of ESOM in intrusion detection for other types of attacks. We also plan to use P-Maps and U*-maps in order to examine if using these maps we may produce better results and provide a comparison between these topology maps and their efficiency in anomaly detection. As future work we will also research the use of this scheme in a distributed solution of the Intrusion Detection problem.

ACKNOWLEDGEMENT

Special thanks to the Databionic research team for the valuable advice and comments concerning the use of the Databionic ESOM tool.
This work has been partially supported by the Greek Research and Technology Secretariat under a PENED grant.

Table 1. Datasets used for evaluation

Datasets	Normal	Total DoS	DoS attacks	
10%KDD	97278	391458	280790 smurf 107201 neptune 979 teardrop	2203 back 264 pod 21 land
DS_1	29126	10697	3695 smurf 5000 neptune	2002 back
DS_2	10517	3000	1000 smurf 1000 neptune	1000 back
DS_3	30180	9816	3695 smurf 4000 neptune 99 teardrop	2002 back 20 pod
DS_4	39238	19938	7001 smurf 397 teardrop	2002 back 17 land
DS_5	39298	13939	5001 smurf 6419 neptune 397 teardrop	2002 back 119 pod 1 land
DS_6	39238	10339	4001 neptune 4001 smurf	2002 back 119 pod 17 land
DS_7	32768	30763	18992 neptune 11258 smurf 394 teardrop	102 pod 17 land
DS_8	10505	9000	5000 neptune	4000 neptune
DS_9	10505	10000	8000 smurf	2000 neptune
DS_10	39320	11950	11950 neptune	
DS_11	10517	4000	4000 neptune	
DS_12	39299	11278	11278 smurf	
DS_13	10518	7563	7563 smurf	
DS_14	15005	15000	15000 neptune	
DS_15	10000	20000	20000 smurf	

Table 2. Evaluation results

Evaluation Experiments	Training Dataset	Testing Dataset	Training parameters	DR	FA
Exp_1	DS_1	DS_2	160x180 neurons 50 epochs	98.9%	0.9%
Exp_ 2	DS_3	DS_2	160x180 neurons 50 epochs	99.46%	1.1%
Exp_3	DS_4	DS_8	160x180 neurons 50 epochs	98.3%	2.9%
Exp_4	DS_5	DS_9	160x180 neurons 50 epochs	99.4%	0.21%
Exp_5	DS_6	DS_8	180x200 neurons 60 epochs	99.17%	0.3%
Exp_6	DS_6	DS_9	180x200 neurons 60 epochs	99.24%	0.32%
Exp_7	DS_7	DS_8	180x200 neurons 60 epochs	98.3%	2.9%
Exp_ 8	DS_10	DS_11	180x200 neurons 60 epochs	98.8%	1.61%
Exp_9	DS_10	DS_14	180x200 neurons	99.43%	0.39%
Exp_10	DS_12	DS_13	160x180 neurons 50 epochs	99.81%	0.1%
Exp_11	DS_12	DS_15	160x180 neurons 50 epochs	99.89%	0.05%
Exp_12	DS_6	DS_15	180x200 60 epochs	99.78%	0.34%

REFERENCES

[1] The Third International Knowledge Discovery and Data Mining Tools Competition, May 2002, Available from http://kdd.ics.uci.edu/databases/kddcup99.kddcup99.html.

[2] S. Haykin, “Neural Networks: A comprehensive Foundation”, Prentice-Hall, New Jersey, USA, 2nd edition, 1999.

[3] L. Girardin, D. Brodbeck, “A Visual Approach for monitoring Logs”, 1998 LISA XII, December 1998, USA.

[4] B.V. Nguyen, “SOM for anomaly detection”, CS680 report, Spring 2002.

[5] A. J. Hoglund, K. Hatonen, A. S. Sorvari, “A computer host-based user anomaly detection system using the SOM”, Proc. of IEEE IJCNN 2000, Vol. 5, pp. 411-416.

[6] P. Lichodzijewski, A.N. Zincir-Heywood, M. I. Heywood, “Host-Based Intrusion Detection using Self-Organizing Maps”, Proc. of IJCNN ’02.

[7] Khaled Labib, V. Rao Vemuri, “NSOM: A Real-time Network-Based Intrusion Detection system Using Self-Organizing Maps”, 2002, Networks security.

[8] B. Rhodes, J. Mahaffey, J. Cannady, “Multiple SOMs for Intrusion Detection”, Proc. of the NISSC 2000 conference.

[9]C. Jirapummin, N. Wattanapongsakorn, P. Kanhamanon, “Hybrid Neural Networks for Intrusion Detection System”, Proc. of ITC – CSCC, pp. 928-931, Thailand, July 2002.

[10] F. Gonzalez, D. Dasgupta, “Neuro-Immune and Self-Organizing Map Approaches to Anomaly Detection”, In Proc. of 1st ICAIS, pp. 203-211, UK, 9-11 September 2002.

[11] G.H. Kayacik, A.N. Zincir-Heywood, M.I. Heywood, "On the capability of SOM based intrusion detection systems," Proc. of IEEE IJCNN, Portland, USA, July 2003.

[12] A. Ultsch, “Data Mining and Knowledge Discovery with Emergent SOFMs for Multivariate Time Series”, In Kohonen Maps, (1999), pp. 33-46.

[13] A. Ultsch, “Maps for visualization of high-dimensional Data Spaces”, Proc. WSOM, Kyushu, Japan, (2003), pp. 225-230.

[14] C. Douligeris, A. Mitrokotsa, “DDoS attacks and defense mechanisms: classification and state-of-the-art”, Computer Networks, Vol. 44, (5), April 2004, pp.643-666

[15] A. Ultsch, F. Moerchen “ESOM-Maps: tools for clustering, visualization, and classification with Emergent SOM”, Tech. Report Dept. of Mathematics and Computer Science, University of Marburg, Germany, (46), (2005).

[16] Databionic ESOM Tools, Available from <http://databionic-esom.sourceforge.net/devel.html>

[17] S. Mukkamala, A. H. Sung, “Identifying significant Features for Network Forensic Analysis Using Artificial Intelligent Techniques”, International Journal of Digital Evidence, Winter 2003, Vol. 1, Issue 4.