IDS Based on Bio-Inspired Models

Paolo Gastaldo[(1)], Francesco Picasso[(1)], Rodolfo Zunino[(1)],
Álvaro Herrero[(2)], Emilio Corchado[(2)] and José Manuel Sáiz[(2)]

[(1)]Dept. of Biophysical and Electronic Engineering (DIBE), Genoa University
Via Opera Pia 11a, 16145 Genoa, Italy
{paolo.gastaldo, francesco.picasso, rodolfo.zunino}@unige.it

[(2)]Department of Civil Engineering, University of Burgos
C/ Francisco de Vitoria s/n, 09006 Burgos, Spain
{ahcosio, escorchado, jmsaiz}@ubu.es

Abstract. Unsupervised projection approaches can support Intrusion Detection Systems for computer network security. The involved technologies assist a network manager in detecting anomalies and potential threats by an intuitive display of the progression of network traffic. Projection methods operate as smart compression tools and map raw, high-dimensional traffic data into 2-D or 3-D spaces for subsequent graphical display. The paper compares three projection methods, namely, Cooperative Maximum Likelihood Hebbian Learning, Auto-Associative Back-Propagation networks and Principal Component Analysis. Empirical tests on anomalous situations related to the Simple Network Management Protocol (SNMP) confirm the validity of the projection-based approach. One of these anomalous situations (the SNMP community search) is faced by these projection models for the first time. This work also highlights the importance of the time-information dependence in the identification of anomalous situations in the case of the applied methods.

Keywords: Unsupervised Learning, Projection Methods, Auto-Associative Back-Propagation, Computer Network Security, Intrusion Detection, Visualization.

1 Introduction

Intrusion Detection Systems (IDS's) [1] monitor traffic in computer networks and take, or suggest, defensive actions when they detect suspect activities. IDS's are common elements in modern infrastructures to enforce network policies. Today's commercial systems typically rely on a knowledge base of rules to discriminate normal from malicious traffic. The set of rules, however, is susceptible to inconsistencies, and continuous updating is required to cover previously unseen attack patterns.

Human beings seem to be able to learn without explicit supervision. One aim of unsupervised learning is to mimic this aspect of human learning and hence this type of learning tends to use more biologically plausible methods than those using error descent methods. Both projection and unsupervised methods [2], [3], [4], [5] can

successfully apply to the visual based IDS's technology. This paper analyzes some of those technologies, and compares the performances of two actual implementations of such paradigms: Cooperative Maximum Likelihood Hebbian Learning (CMLHL) [6] and Auto-Associative Back-Propagation (AABP) neural networks [7].

2 Visual Inspection of Traffic in Modern IDS's

A visual-based IDS could be organized as follows (Fig. 1):

- packets traveling across the network are intercepted by a capture device;
- traffic is coded by a set of features spanning a multidimensional vector space;
- the compression component operates on feature vectors and yields a two or three-dimensional representation of the network traffic;
- the outcome of the projection component are presented to the network manager on a display device.

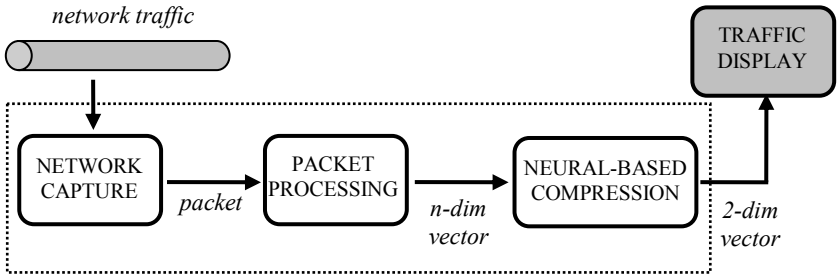

Fig. 1. A general schema of a visual-based Intrusion Detection System.

The compression component processes an *n*-dimensional vector that has been previously assembled by a "packet processing" module, which extracts numerical features associated with each network packet. Hence, unlike other models which operate at the connection level [8], [9], the proposed IDS operates at the packet level. The design of the feature set is a crucial issue that has been thoroughly addressed in the literature [10]. It has been proved that timestamp, source and address port, and protocol uniquely identify a connection [11]. When dealing with Transmission Control Protocol (TCP) traffic, additional features may be required (e.g. to track connection state [12]); instead, User Datagram Protocol (UDP) traffic can be effectively characterized by a reduced feature set [4].

The compression component clearly is the core of the overall IDS, and is designed to yield an effective representation of network traffic. Its overall function is to support the network supervisor at detecting traffic anomalies. A connectionist approach seems to best fit the anomaly-detection problem mainly because it allows a system to learn empirically the input-output relationship between raw traffic data and its subsequent interpretation. Indeed, the outlier-detection method does not require any a-priori analytical formulation of the underlying phenomenon. Section 3 illustrates two different projection approaches to the dimensionality reduction mapping task of traffic data.

3 Nonlinear Projection Methods for Dimensionality Reduction

3.1 Neural Implementation of Exploratory Projection Pursuit

The classical statistical method of Exploratory Projection Pursuit (EPP) [3] provides a linear projection of a data set onto a set of basis vectors which best reveal the interesting structure in data; interestingness is usually defined in terms of how far the distribution is from the Gaussian distribution. Maximum Likelihood Hebbian Learning (MLHL) [13] is a neural implementation of EPP: it identifies interestingness by maximising the probability of the residuals under specific probability density functions which are non-Gaussian.

The Cooperative Maximum Likelihood Hebbian Learning (CMLHL) model [6] extends the MLHL paradigm by adding lateral connections [6], which have been derived from the Rectified Gaussian Distribution [14]. The resultant net can find the independent factors of a data set but does so in a way that captures some type of global ordering in the data set. Considering a D-dimensional input vector ($\mathbf{x}$), and a Q-dimensional output vector ($\mathbf{y}$), with W_{ij} being the weight (linking input j to output i), then CMLHL can be expressed [6] as:

1. Feed-forward step:

$$y_i = \sum_{j=1}^{D} W_{ij} x_j \; ; \; i = 1,...,Q \cdot \tag{1}$$

2. Lateral activation passing:

$$y_i(t+1) = \left[y_i(t) + \tau\left(b - Ay\right) \right]^{+} ; \; i = 1,...,Q \, . \tag{2}$$

3. Feedback step:

$$e_j = x_j - \sum_{i=1}^{Q} W_{ij} y_i \; ; \; j = 1,...,D \cdot \tag{3}$$

4. Weight change:

$$\Delta W_{ij} = \eta . y_i . sign\left(e_j\right) | \, e_j \, |^{p-1} ; \; i = 1,...,Q; \; j = 1,...,D \, . \tag{4}$$

Where: η is the learning rate, τ is the "strength" of the lateral connections, b is a bias parameter, p a parameter related to the energy function [6] and A a symmetric matrix used to modify the response to the data [6]. The effect of this matrix is based on the relation between the distances separating the output neurons.

3.2 Auto-Associative Back-Propagation Networks

An Auto-Associative Back-Propagation (AABP) network involves a Multi-Layer Peceptron (MLP) structure [15] which typically includes at least one "hidden" layer. It associates an input vector, $\boldsymbol{x} \in \Re^D$, with an output vector, $\boldsymbol{y} \in \Re^Q$, computed as:

$$y_q(\mathbf{x}) = w'_{q,0} + \sum_{u=1}^{N_h} \left[w'_{u,q} \cdot \sigma\left(w_{u,0} + \sum_{k=1}^{D} w_{u,k} x_k \right) \right] ; \; q = 1,...,Q \cdot \tag{5}$$

where, N_h is the depth of the sigmoid series expansion, and W represents the coefficients of the weights for the interconnections between the two upper layers. Empirical fitting drives the weights, W, so that the network best reproduces (in an Euclidean sense) the desired ($\boldsymbol{x}$, $\boldsymbol{y} \equiv \boldsymbol{x}$) mapping over a given training set.

Auto-associativity in the AABP model stems from imposing the output target equal to the input stimulus, hence $\boldsymbol{y} \equiv \boldsymbol{x}$ and $Q = D$; compression instead results from imposing that $N_h < D$. The overall goal of this unsupervised approach is a reduction in dimensionality by forcing the network to replicate the training sample distribution. At run-time, an AABP network associates each input vector with the "coding" values computed by the hidden neurons. Since the hidden layer is typically smaller than the input/output ones, these mapping outputs support the (lossy) transformation from the input space into a lower-dimensional representation. A three-layer AABP network implements a mapping that is, in fact, affine to linear Principal Component Analysis (PCA).

Non-Linear Principal Component Analysis (NLPCA) [7] was designed to circumvent the limitations of linearity inherent in the PCA model (Fig. 2). The output layer still imposes the auto-associative constraint $\boldsymbol{y} \equiv \boldsymbol{x}$, and a hidden layer supports dimensionality reduction. Both the compression and reconstruction sections, however, include an additional layer of neurons, thus yielding a five-layer network. At run-time after completion of training, the mapping outputs of the middle "coding" layer provide the low-dimensional representation of each input vector. The increase in power of representation conveyed by the NLPCA augmentation is remarkable. Typically, the "reconstruction" section and the compression section will always be symmetrical and will therefore always yield equivalent, universal (inverse) mapping capabilities.

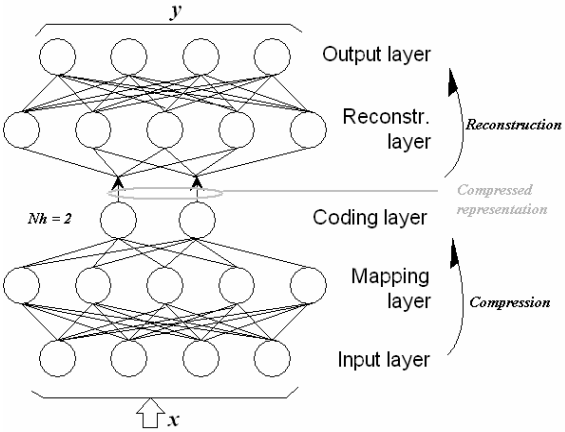

Fig. 2. The Non-Linear AABP network.

The main advantage is that the compressed representation does not relate to any linear model but stems from a mostly general representation that is learned empirically. NLPCA techniques fit those domains in which 1) a non-linear representation best encompasses the observed empirical phenomenon, and at the same time, 2) a considerable number of empirical samples are available.

4 Using Projection Methods for Visual Based IDS's

In principle, any unsupervised method applies to the involved representation process; Self-Organizing Maps [1] and Vector Quantization-based methods [5] have had a considerable success in supporting IDS's technology. As compared with those models, projection based approaches offer the crucial advantage of combining compression ability and the support for graphical, intuitive representation. The following sections discuss the performance of the CMLHL model and the AABP model in the monitoring and detection of anomalous events in a real scenario.

4.1 Identification of SNMP Community Search

For simplicity and without loss of generality, the present IDS is targeted to detect traffic anomalies within the Simple Network Management Protocol (SNMP), which is a part of the Transmission Control Protocol/Internet Protocol (TCP/IP) protocol suite. SNMP is an application layer protocol that supports the exchange of management information between network devices. This protocol enables network administrators to manage network performance and is used to control network elements such as routers, bridges and switches. This property makes SNMP data quite sensitive and liable to potential attacks. Indeed, an attack based on the SNMP may severely compromise system security, as reported by CISCO [16].

In order to test the above mentioned models under the IDS's field, this research addresses some different types of attacks relying on the SNMP:

- SNMP port sweep: a port scan (or sweep) is an attempt to count the services running on a machine (or a set of machines) by probing each port for a response.
- SNMP community search: the community string can be seen as the SNMP password for versions 1 and 2. The intruder sends SNMP queries to try and determine the SNMP community string.
- MIB information transfer: the Management Information Base (MIB) collects information on each managed device, including sensitive data such as network and router information (e.g. IP and MAC addresses and VLan configuration). As specified by the Internet Activities Board, the SNMP is used to access MIB objects; thus, protecting a network from malicious MIB information transfer is crucial.

Two different data sets have been used in the present work:

- Data set 1 (utilized in [4]): it contains three consecutive sweeps on some machines. Port numbers 161, 162 and 3750 have been probed for SNMP. This is followed by an MIB information transfer (for a known or default community name).
- Data set 2: it differs from the previous one in that the attacker probes for 3 different community names on 3 different ports. That is, 3 different hosts try to check 3 different community strings ("public", "private" and "aab") for SNMP on ports 161, 1161 and 2161 at the same time. This attack is set between the port sweeps and the MIB transfer. After trying to find whether and where (hosts and port numbers) SNMP is running, the community string is requested in order to access the information contained in the MIB. When the community string has been found,

the MIB information is read. It is worth highlighting that the projection models face the SNMP community search (included in this data set) for the first time.

Both data sets contained network packets related to normal and anomalous situations. They were captured from UDP traffic, as SNMP uses UDP as the transport protocol. Hence, the data sets included only packets that use UDP as transport layer, IP as network layer and some protocols (SNMP among them) as application layer.

4.2 Feature Extraction

The eventual network-based IDS for the detection of SNMP anomalous traffic is structured as shown in Fig. 1. The "packet processing" component generates feature vectors by working out information contained in the packet header. In the present research, network packets are characterized by using the set of features that already proved to be effective for detection of anomalous SNMP traffic [4]. The set of five features that are extracted from packets contribute to build up the neural-network input vector, $\mathbf{x} \in \Re^5$. These features can be listed as follows: protocol ID (the protocol of the packet), source port (the port number of the device that sent the packet), destination port (the port number of the host to which the packet was sent), size (packet size in Bytes) and time (the time when the packet was sent).

As such, at the output of the "packet processing" module the network traffic is mapped in a five-dimensional feature space. According to the set-up discussed in Section 2, the compression module exploits the projection method (either CMLHL or AABP) to generate a two-dimensional representation of the network traffic by starting from the original five-dimensional space. Thus, first an offline training phase uses empirical data to set the configuration of weight quantities for the neural network. Then, the eventual neural system is used to process the feature vectors generated at run-time and to feed the visual display.

4.3 Identifying Different Anomalous Situations

The following experimental verification shows the performance of three projection models (linear PCA, CMLHL and AABP) to analyse their response to the data sets described above. In these experiments, the configuration of the AABP network included a number of 30 nodes in the hidden layers (coding and reconstruction), while of course the number of neurons in the middle layer was N_h=2.

As can be seen in Fig. 3.a, the CMLHL projections manage to identify all the anomalous situations contained in Data set 2. The three sweeps (Group A) and the community search (Group B) are identified due to their non-parallel direction to the normal one. The MIB information transfer (Groups C and D) is identified due to its orthogonal direction with respect to the normal traffic (vertical and parallel straight lines) and to the high density of packets. Linear PCA projection was tested on the same data for comparison (Fig. 3.b). PCA could identify the port sweeps (Group A) and the community search (Group B), but failed at detecting the MIB information transfer: in the graph the MIB packets (Groups C and D) evolve in a direction parallel to the "normal" one.

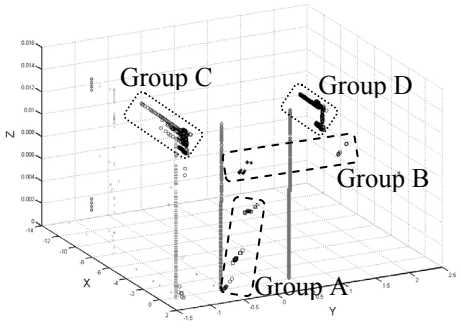

Fig. 3.a tridimensional CMLHL projection

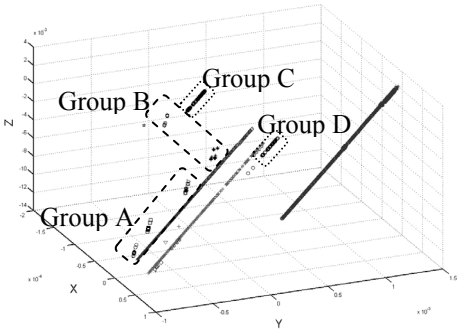

Fig. 3.b tridimensional PCA projection

Fig. 3. Experimental verification of the projection methods for Data set 2: a) 3-D CMLHL projection; b) 3-D PCA projection.

It can be seen in Fig. 4.a how CMLHL is able to identify the anomalous situations (Groups A, B and C) included in Data set 1 as it has been explained before. This outcome shows the intrinsic robustness of CMLHL, which is able to properly respond to a complex data set that includes time as a variable. In the case of AABP, this issue is not clear at all when using the time as one of the input variables. Including a linear growing variable (time) forced the two inner neurons to learn a linear path, which is reflected in Fig. 4.b.

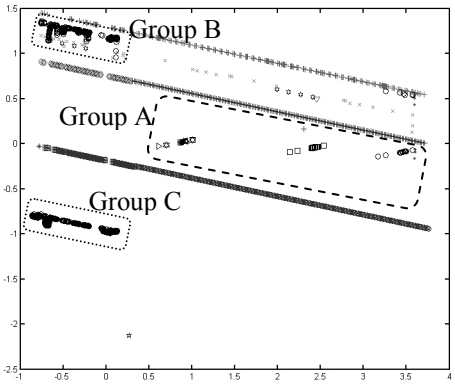

Fig. 4.a bidimensional CMLHL projection

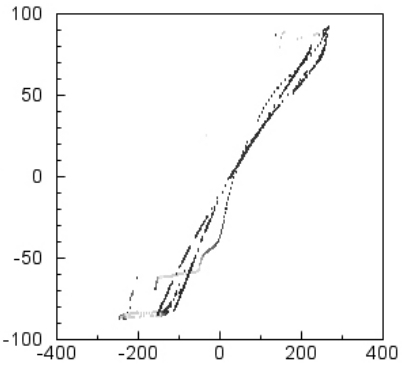

Fig. 4.b bidimensional AABP projection

Fig. 4. Bidimensional projections for Data set 1 (including time in the input variables): a) CMLHL projection; b) AABP projection.

The comparative approach supports the view that projection methods are an interesting and powerful tool in the identification of anomalous situations. This study shows for the first time how the CMLHL is capable of identifying the anomalous situation related to an SNMP community search. Another interesting issue is the capability of CMLHL to process time information as one of the data variables. In contrast, the AABP network can not deal with the time variable, even though it obtains similar results to CMLHL by plotting its results against time [17]. The model

can be extended by the application of multiagent systems in an architecture akin to the MObile VIsualization Connectionist Agent-Based IDS (MOVICAB-IDS) [18].

Acknowledgments. This research has been partially supported by the MCyT project TIN2004-07033.

References

1. Laskov, P., Dussel, P., Schafer, C., Rieck, K.: Learning Intrusion Detection: Supervised or Unsupervised? In: ICIAP 2005. LNCS, vol. 3617, pp. 50-57. Springer, Heidelberg (2005)
2. Hertz, J.A., Krogh, A., Palmer, R.G.: Introduction to the Theory of Neural Computation. Addison-Wesley, Redwood City, CA (1991)
3. Friedman, J.H., Tukey, J.W.: A Projection Pursuit Algorithm for Exploratory Data-Analysis. IEEE Transactions on Computers 23(9), 881-890 (1974)
4. Corchado, E., Herrero, A., Sáiz, J.M.: Detecting Compounded Anomalous SNMP Situations Using Cooperative Unsupervised Pattern Recognition. In: ICANN 2005. LNCS, vol. 3697(2), pp. 905-910. Springer, Heidelberg (2005)
5. Zheng, J., Hu, M.: An Anomaly Intrusion Detection System Based on Vector Quantization. IEICE - Trans. Inf. Syst. E89-D(1), 201-210 (2006)
6. Corchado, E., Fyfe, C.: Connectionist Techniques for the Identification and Suppression of Interfering Underlying Factors. Int J Patt Recogn Artif Intell 17(8), 1447-1466 (2003)
7. Kramer, M.A.: Nonlinear Principal Component Analysis Using Autoassociative Neural Networks. AIChE Journal 37(2), 233-243 (1991)
8. Lippmann, R., Haines, J.W., Fried, D.J., Korba, J., Das, K.: Analysis and Results of the 1999 DARPA Off-Line Intrusion Detection Evaluation. In: RAID 2000. LNCS, vol. 1907, pp. 162 - 182. Springer, Heidelberg (2000)
9. Sabhnani, M., Serpen, G: Application of Machine Learning Algorithms to KDD Intrusion Detection Dataset within Misuse Detection Context. In: 2003 Int. Conference on Machine Learning, Models, Technologies and Applications. pp. 623-630 (2003)
10. Lee, W., Xiang, D.: Information-theoretic Measures for Anomaly Detection. In: Proc. IEEE Symposium on Security and Privacy (S&P 2001). pp. 130 - 143 (2001)
11. Lee, W., Stolfo, S.J., Mok, K.W.: Mining in a Data-Flow Environment: Experience in Network Intrusion Detection. Proc. 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM Press, San Diego, California, USA (1999)
12. Lee, W., Stolfo, S.J., Mok, K.W.: Adaptive Intrusion Detection: A Data Mining Approach. Artificial Intelligence Review 14(6), 533-567 (2000)
13. Corchado, E., MacDonald, D., Fyfe, C.: Maximum and Minimum Likelihood Hebbian Learning for Exploratory Projection Pursuit. Data Mining and Knowledge Discovery 8(3), 203-225 (2004)
14. Seung, H.S., Socci, N.D., Lee, D.: The Rectified Gaussian Distribution. Advances in Neural Information Processing Systems 10, 350-356 (1998)
15. Rumelhart, D.E., McClelland, J.L.: Parallel Distributed Processing. MIT Press, MA (1986)
16. Cisco Secure Consulting: Vulnerability Statistics Report. 2000
17. Herrero, A., Corchado, E., Gastaldo, P., Zunino, R.: A Comparison of Neural Projection Techniques Applied to Intrusion Detection Systems. Proc. 9th International Work-Conference on Artificial Neural Networks - IWANN'2007. LNCS, vol. 4507, pp. 1138-1146. Springer, Heidelberg (2007) (*"In press"*)
18. Herrero, A., Corchado, E., Sáiz, J.M.: MOVICAB-IDS: Visual Analysis of Network Traffic Data Streams for Intrusion Detection. In: Corchado, E., Yin, H., Botti, V., Fyfe, C. (eds.) IDEAL 2006. LNCS, vol. 4224, pp. 1424-1433. Springer, Heidelberg (2006)