Identification of Anomalous SNMP Situations Using a Cooperative Connectionist Exploratory Projection Pursuit Model

Álvaro Herrero, Emilio Corchado, José Manuel Sáiz

Department of Civil Engineering, University of Burgos, Spain
escorchado@ubu.es

Abstract. The work presented in this paper shows the capability of a connectionist model, based on a statistical technique called Exploratory Projection Pursuit (EPP), to identify anomalous situations related to the traffic which travels along a computer network. The main novelty of this research resides on the fact that the connectionist architecture used here has never been applied to the field of IDS (Intrusion Detection Systems) and network security. The IDS presented is used as a method to investigate the traffic which travels along the analysed network, detecting SNMP (Simple Network Management Protocol) anomalous traffic patterns. In this paper we have focused our attention on the study of two interesting and dangerous anomalous situations: a port sweep and a MIB (Management Information Base) information transfer. The presented IDS is a useful visualization tool for network administrators to study anomalous situations related to SNMP and decide if they are intrusions or not. To show the power of the method, we illustrate our research by using real intrusion detection scenario specific data sets.

1 Introduction

Connectionist models have been identified as a very promising method of addressing the intrusion detection problem due to two main features: they are suitable to detect day-0 attacks (new and modified intrusion strategies) and they have the ability to classify patterns (attack classification, alert validation). IDS are hardware or software systems that monitor the events occurring in a computer system or network, analyzing them to identify computer security problems in an automate way. IDS have become a necessary additional tool to the security infrastructure of most organizations as the number of network attacks has increased very fast during the last years.

IDS try to identify any attack that may compromise the integrity, confidentiality or availability of a system, which are the three computer security principles [1]. Intrusions are produced for example by attackers accessing to the system from networks as Internet, for authorized users who attempt to obtain more privileges for which they are not authorized and authorized users who misuse the privileges given to them. The complexity increases in the case of distributed network-based systems and insecure networks.

Up to now, there have been several attends to apply artificial neural architectures (such as Self Organising Maps [2, 3], Elman Network [4]) to the network security field [5, 6]. This paper presents an IDS based on a novel neural EPP architecture.

EPP [7, 8, 9, 10] is designed for analyzing high-dimensional data using low-dimensional projections. The aim of EPP is to reveal possible interesting structures hidden in the high-dimensional data so that a human can investigate the projections by eye. This technique can be very interesting for a network administrator to visualize the traffic travelling along the network and use it to detect anomalous situations.

The remainder of this paper is structured as follows:

Section 2 introduces the developed IDS model. Section 3 provides an overview of the unsupervised connectionist method used by the IDS model. Section 4 describes the problem and Section 5 describes the data set used. Finally, we present the results in Section 6 and the conclusions and future research in Section 7.

2 A Novel IDS Model

The aim of this work is the design of a system capable of detecting anomalous situations for a computer network. The information analysed by our system is obtained from the packets which travel along the network. So, it is a Network-Based IDS. The necessary data for the traffic analysis is contained on the captured packets headers. This information can be obtained using a network analyser.

When we talk about anomaly detection models we refer to IDS which detect intrusions by looking for abnormal network traffic. Anomaly detection is based on the assumption that misuse or intrusive behaviour deviates from normal system use [5, 11]. In many cases, as in the case of the attacker who breaks into a legitimate user's account, this is a right assumption.

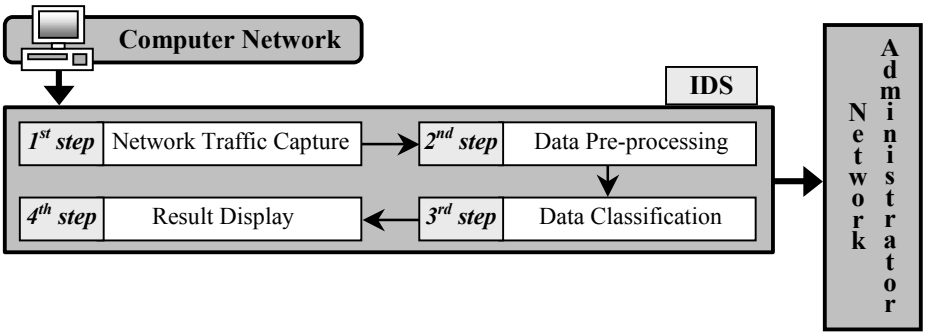

Fig. 1. Graphical Description of the proposed IDS

The structure of this novel IDS is showed in Fig.1 and it is described as follows:

- **1st step**.- Network Traffic Capture: one of the network interfaces is set up as "promiscuous" mode. It captures all the packets travelling along the network.
- **2nd step**.- Data Pre-processing: the captured data is pre-processed as it is described in section 5 and later, it is used as the input data to the following stage.

– **3rd step**.- Data Classification: once the data has been pre-processed, the connectionist model (section 3) analyses the data and identifies the anomalous patterns.
– **4th step**.- Result Display: the last step is related to the visualization stage. At the end, the output of the network is presented to the administrator or person in charge of the network security.

3 The Unsupervised Connectionist Architecture

EPP [7, 8, 9, 10] is a statistical method for solving the difficult problem of identifying structure in high dimensional data. The method used here is based on the projection of the data onto a lower dimensional subspace in which we search for its structure by eye. It is necessary to define an "index" that measures the interestingness of a projection. After that, the data is transformed by maximizing the index in order to maximize the interest according to that index. From a statistical point of view the most interesting directions are those which are as non-Gaussian as possible.

The Data Classification and Result Display steps (Fig. 1) performed by this IDS model are based on the use of a neural EPP model called Cooperative Maximum Likelihood Hebbian Learning (CMLHL) [12, 13, 14]. It was initially applied to the field of Artificial Vision [12, 13] to identify local filters in space and time. Here, we have applied it to the computer security field [6]. It is based on Maximum Likelihood Hebbian Learning (MLHL) [9, 10, 13]. Consider a N-dimensional input vector, x, and a M-dimensional output vector, y, with W_{ij} being the weight linking input j to output i.

MLHL can be expressed as:

$$y_i = \sum_{j=1}^{\mathbf{N}} W_{ij} x_j, \forall i \,. \tag{1}$$

The activation is fed back through the same weights and subtracted from the input.

$$e_j = x_j - \sum_{i=1}^{M} W_{ij} y_i, \forall j \,. \tag{2}$$

And finally, the weight update:

$$\Delta W_{ij} = \eta . y_i . sign\left(e_j\right) | e_j |^{p-1} \,. \tag{3}$$

Lateral connections [12, 13] have been derived from the Rectified Gaussian Distribution [15] and applied to the negative feedback network [16]. The resultant net will be shown to be a network which can find the independent factors of a data set but do so in a way which captures some type of global ordering in the data set.

We use the standard MLHL but now with lateral connections (which act after the feed forward but before the feedback).

Thus we have a feed forward step (Eq. 1) follows by:

$$\text{Lateral activation passing: } y_i(t+1) = \left[y_i(t) + \tau\left(b - Ay\right)\right]^+. \quad (4)$$

$$\text{Feedback: } e_j = x_j - \sum_{i=1}^{M} W_{ij} y_i \text{ . (Eq. 2)} \quad (5)$$

$$\text{Weight change: } \Delta W_{ij} = \eta . y_i . sign\left(e_j\right) | e_j |^{p-1} \text{ . (Eq. 3)} \quad (6)$$

Where:

- τ is the "strength" of the lateral connections.
- b is the bias parameter.
- A is a symmetric matrix used to modify the response to the data based on the relation between the distances among the output neurons.
- η is the learning rate.
- p is a parameter related to the energy function [9, 10, 13].

4 Problem Description

A protocol in a network context is a specification that describes low-level details of host-to-host interfaces or high-level exchanges between application programs. Among all the implemented network protocols, there are several of them that can be considered quite more dangerous (in terms of network security), such as SNMP, ICMP (Internet Control Message Protocol), TFTP (Trivial File Transfer Protocol) and so on.

We have focused our effort in the study of SNMP anomalous situations because an attack based on this protocol may severely compromise the systems security. CISCO [17] found the top five most vulnerable services in order of importance, and SNMP was one of them. Initially, SNMP was oriented to manage nodes in the Internet community [18]. It is used to read and write a wide variety of information about the device: operating system, version, routing tables, and so on. Some of this data can be extremely sensitive and the MIB is used to store this information. The MIB can be roughly defined as a database which contains information about some elements or devices that can be network-controlled.

One special feature of the traffic travelling along the network is that SNMP packets are generated and sent inside the own network. It is an internal protocol and any host out of the network can not introduce any packets of this type in the network.

There are some anomalous situations related to SNMP implementations. Among those we have chosen the two most dangerous ones: a SNMP port sweep and a MIB information transfer. This kind of transfer is considered a quite dangerous situation because a person having some free tools, some basic SNMP knowledge and the community password (in SNMP v. 1 and SNMP v. 2) can come up with all sorts of interesting and sometimes useful information.

The study of SNMP is the reason why the system selects packets based on UDP (User Datagram Protocol) in the Data Pre-processing step. This means that in terms of the TCP/IP (Transmission Control Protocol/Internet Protocol) protocol stack, the model captures only the packets using UDP at transport layer and IP at network layer.

5 Real Intrusion Detection Scenario Specific Dataset

In the Data Pre-processing step, the system performs a data selection of all the information captured: we used the following 5 variables extracted from the packet headers:

- **Timestamp**: time (in milliseconds) when the packet was sent (difference in relation to the first captured packet).
- **Protocol**: we have codified all the protocols contained in the data set, taking values between 1 and 40.
- **Source Port**: port number of the source host which sent the packet.
- **Destination Port**: port number of the destination host where the packet was sent.
- **Size**: total packet size (in bytes).

As it is said before, we study two anomalous situations related to SNMP. Therefore the used data set contains examples of those situations:

- **SNMP port sweep**: it consists of a scanning of network hosts for the SNMP port using sniffing methods. The aim is to make a systematic sweep in a group of hosts to verify if SNMP is active in some ports. We have used default [18] port numbers (161 and 162) and also a random port number (3750) as a test random element.
- **MIB information transfer**: the previous situation is followed by a MIB transfer. The data set contains a transfer of some information stored in the SNMP MIB.

In addition to the SNMP packets, this data set contains traffic related to other protocols installed in the network, like NETBIOS and BOOTPS.

6 Results

Fig. 2 shows the best projection displayed by the model for this data set, where it is easy to identify some anomalous groupings of packets. Just by looking, a strange behaviour can be identified in Groups 1, 2 and 3. After a visual analysis of these groups, the following characteristics can be identified:

- Group 1 (Fig. 2) is subdivided in three subgroups. Packets belonging to each one of these subgroups (S1, S2 and S3) progress in a direction (Abnormal direction) different from the one in which the rest of the packets (”normal”) groups progress (Normal direction).
- Groups 2 and 3 (Fig. 2) have a very high temporal concentration of packets. Moreover, packets contained in these groups do not progress in a unique direction, they progress in two different ones, while the rest of the packets (”normal”) groups progress in only one direction (Normal direction).

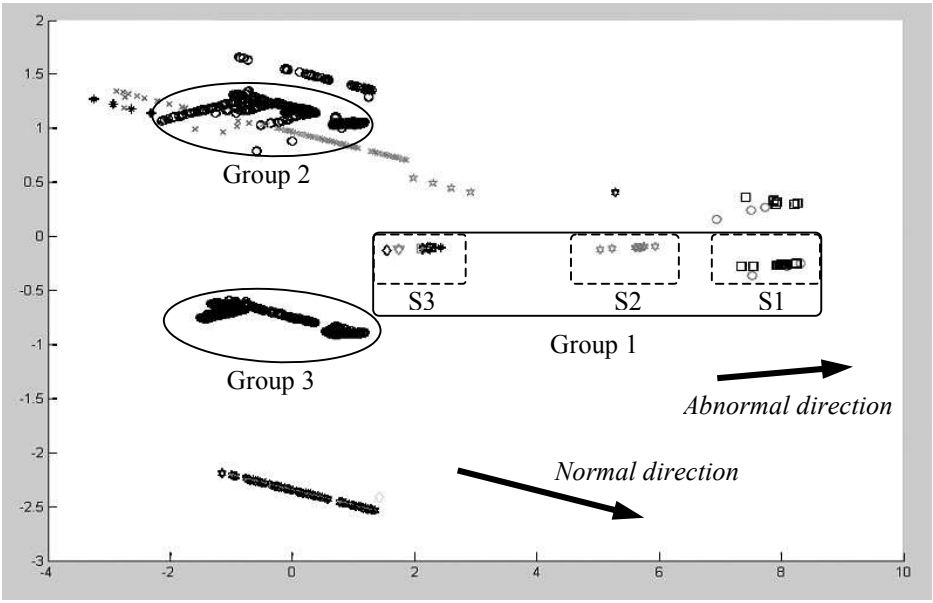

Fig. 2. Data projection displayed by the model

After the study of these graphical characteristics and the analysed data set, we have came to the following conclusions:

- Group 1 (Fig. 2) contains packets related to the SNMP sweep presented in the data set. Each one of the subgroups in it are associated to each port number included in the sweep (161, 162 and 3750 in this case). This is, the first subgroup starting from the right side (S1) includes packets sent to port number 161, the second subgroup (S2) includes packets sent to port number 162 and so on.
- Groups 2 and 3 (Fig. 2) are related to the SNMP MIB transfer mentioned above. They contain packets sent and received during the transfer embedded in the data set. Group 2 contains all the traffic in one way (from destination to source), while Group 3 contains all the traffic in the other way (from source to destination). One of the main axes identified in the visual analysis can be related to the packet size and the protocol while the other one is related to the timestamp.

We have compared the results provided by CMLHL (Fig. 2) with MLHL (Fig.3.a) and Principal Component Analysis (PCA) (Fig.3.b). CMLHL is able of identifying both anomalous situations while PCA [16, 19] is just identifying the sweep (Group 1, Group 2 and Group 3 in Fig.3.b) by means of normal/abnormal directions. Fig.3.a shows how MLHL is capable of detecting the MIB transfer (Groups 2 and 3) but the port sweep (Group 1) is not detected as clearly as by using CMLHL.

CMLHL highlights anomalous situation more clearly because the projections are more spread out so it is easier to analyse them. The anomalous situations are detected due to the different traffic directions or the high temporal density. CMLHL shows both features better than other methods as MLHL or PCA.

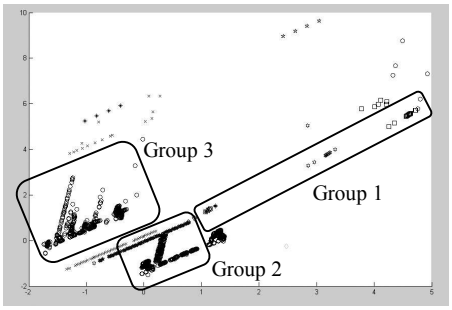

Fig. 3.a. MLHL Projection

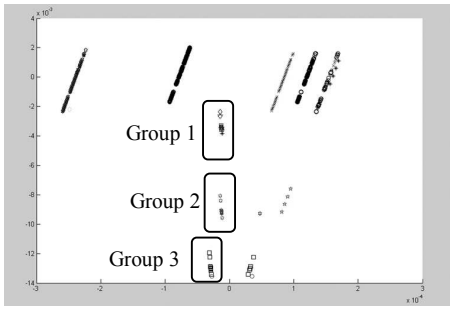

Fig. 3.b. PCA Projection

7 Conclusions and Future Work

The visualization tool used in the Result Display step, displays data projections highlighting anomalous situations clearly enough to alert the network administrator, taking into account aspects as the traffic density or "anomalous" traffic directions.

It has been showed how the proposed model is able to identify anomalous situations by means of the eyes of the network administrator. In this case, the system has identified a port sweep followed by a MIB transfer. Both anomalous situations contained in the data set have been identified by the model. So in performance terms, the systems achieves very good results because it is able to detect all the anomalous situations contained in the data set.

We can not consider a simple packet as an anomalous one, because it is considered anomalous with respect to the rest of packets (both normal and anomalous ones) traveling along the network.

Further work will be focused on:

- Application of GRID [20, 21] computation. This increase of the system power will be used in such a way that the system can be able to capture, process, classify and display the data in real time.
- Study of different distributions and learning rules.
- Application to more complex data sets, trying to extend the model to cover several different situations, including other SNMP anomalous situations and protocols, until to cover all of them.

Acknowledgments

This research has been supported by the McyT projects: TIN2004-07033.

References

1. Myerson, J.M.: Identifying Enterprise Network Vulnerabilities. International Journal of Network Management, Vol. 12 (3) (2002) 135–144
2. Hätönen, K., Höglund, A.,Sorvari, A.: A Computer Host-Based User Anomaly Detection System Using the Self-Organizing Map. International Joint Conference of Neural Networks, Vol. 5 (2000) 411–416
3. Zanero S., Savaresi S.M.: Unsupervised Learning Techniques for an Intrusion Detection System. ACM Symposium on Applied Computing (2004) 412-419
4. Ghosh, A. Schwartzbard A., Schatz A.: Learning Program Behavior Profiles for Intrusion Detection. Workshop on Intrusion Detection and Network Monitoring (1999) 51-62
5. Debar, H., Becker, M., Siboni, D.: A Neural Network Component for an Intrusion Detection System. IEEE Symposium on Research in Computer Security and Privacy. Oakland, California (1992)
6. Corchado, E., Herrero, A., Baruque, B., Saiz J.M.: Intrusion Detection System Based on a Cooperative Topology Preserving Method. International Conference on Adaptive and Natural Computing Algorithms. Lecture Notes in Computer Science. Springer-Verlag, Berlin Heidelberg New York (2005) 454-457
7. Friedman J., Tukey. J.: A Projection Pursuit Algorithm for Exploratory Data Analysis. IEEE Transaction on Computers, Vol. 23 (1974) 881-890
8. Hyvärinen A.: Complexity Pursuit: Separating Interesting Components from Time Series. Neural Computation, Vol. 13(4) (2001) 883-898
9. Corchado, E., MacDonald, D., Fyfe, C.: Maximum and Minimum Likelihood Hebbian Learning for Exploratory Projection Pursuit. Data Mining and Knowledge Discovery, Vol. 8(3). Kluwer Academic Publishing, (2004) 203-225
10. Fyfe, C., Corchado, E.: Maximum Likelihood Hebbian Rules. European Symposium on Artificial Neural Networks (2002)
11. Denning, D.: An Intrusion Detection Model. IEEE Transactions on Software Engineering, Vol. SE-13(2) (1987)
12. Corchado, E., Han, Y., Fyfe, C.: Structuring Global Responses of Local Filters Using Lateral Connections. Journal of Experimental and Theoretical Artificial Intelligence, Vol. 15(4) (2003) 473-487
13. Corchado, E., Fyfe, C.: Connectionist Techniques for the Identification and Suppression of Interfering Underlying Factors. International Journal of Pattern Recognition and Artificial Intelligence, Vol. 17(8) (2003) 1447-1466
14. Corchado, E., Corchado, J.M., Sáiz, L., Lara, A.: Constructing a Global and Integral Model of Business Management Using a CBR System. First International Conference on Cooperative Design, Visualization and Engineering (2004) 141-147
15. Seung, H.S., Socci, N.D., Lee, D.: The Rectified Gaussian Distribution. Advances in Neural Information Processing Systems, Vol. 10 (1998) 350-356
16. Fyfe, C.: A Neural Network for PCA and Beyond. Neural Processing Letters, Vol. 6(1-2) (1997) 33-41
17. Cisco Secure Consulting: Vulnerability Statistics Report (2000)
18. Case, J., Fedor, M.S., Schoffstall, M.L., Davin, C.: Simple Network Management (SNMP). RFC-1157 (1990)
19. Oja, E.: Neural Networks, Principal Components and Subspaces. International Journal of Neural Systems, Vol. 1 (1989) 61-68
20. Foster I., Kesselman C.: The Grid: Blueprint for a New Computing Infrastructure. 1st edn. Morgan Kaufmann Publishers (1998)
21. Kenny, S.: Towards a Grid-wide Intrusion Detection System. European Grid Conference. Lecture Notes in Computer Science. Springer-Verlag, Berlin Heidelberg New York (2005)