General Letters in Mathematics (GLM)

Journal Homepage: <http://www.refaad.com/views/GLM/home.aspx>

Contribution of wavelets to cybersecurity: Intrusion detection systems using neural networks

Saiida Lazaar^{a,*}

^aslazaar@uae.ac.ma

Departement of Mathematics and Computer Science - ERMIA TEAM. National School of Applied Sciences - Tangier, ABDELMALEK ESSAADI University, Morocco

Abstract

The gigantic growth of the exchanged digital data has raised important security challenges. In this ecosystem, connected objects, systems and networks are exposed to various cyber threats endangering sensitive data and compromising confidentiality, integrity and authentication. Modelling intrusion detection systems (IDS) constitute an important research field with a major goal to protect targeted systems and networks against malicious activities. Many network IDS have been recently designed with artificial intelligence techniques. Signal processing techniques have been applied in network detection systems due to their ability to help for a good intrusion detection. At the same context, the wavelet transform which is considered as a very efficient tool for the decomposition and reconstruction of signals can be recommended in the design of powerful network detection systems, and can be applied for data preprocessing denoising and extracting information. Wavelets combined to neural networks can be useful for modelling intrusion detection with the main challenges to reduce the false alarms, increase the test accuracy and increase novel attacks detection rate. In this work, we present a major contribution in the research field to better understand how wavelets and neural networks can be combined for modelling efficient IDS.

Keywords: Intrusion detection systems, Network Security, Wavelet.

2010 MSC: MSC code1, MSC code2, more.

1. Introduction

The world is experiencing an important growth of digital data exchanged through a multitude of networks spread across the globe. However, these data are often exposed to risks of passive or active attacks. The need to secure the sensitive data is highly expressed, and the protection mechanisms must evolve with the permanent evolution of the attacks. The most recommended solutions deployed to secure data, systems and networks include, in addition to encryption techniques, anti-virus, firewalls, network secure protocols (SSL, IPSEC, VPN, MPLS), intrusion and prevention detection systems. To protect the cyber space against cyber-attacks, IDS are highly stressed for monitoring hosts, monitoring the behaviour of targeted networks, analysing the captured traffic, and identifying malicious or unauthorised activities. Different IDS (signature-based IDS, Network-based IDS, etc.) using different techniques have been proposed about ten years ago [1, 2]. Recently, the scientific research has applied signal processing to identify anomalies, to study characteristics of networks, and to build detection models able to identify

*Corresponding author

Email address: slazaar@uae.ac.ma (Saiida Lazaar)

doi:[10.31559/glm2021.10.2.2](https://doi.org/10.31559/glm2021.10.2.2)

Received 30 Apr 2021 : Revised : 25 May 2021 Accepted: 25 Jun 2021

and to detect various sets of network anomalies with high detection rate, and low false alarm rate. The design of IDSs often starts from signal processing for denoising information. Different techniques exist for the preprocessing step; and the most attractive and promising one uses wavelet transforms. We note that there exists an important scientific contribution dealing with wavelets applied to perform IDS for protecting networks against intrusions and external attacks. Wavelet transforms have been applied as a robust tool to represent the Internet traffic and to improve the efficiency of intrusion detection systems when coupled to neural networks techniques. In the same context, spectral analysis have been used to represent the traffic data in the frequency domain; this traffic is first a set of TCP/IP packets. The main goal in using the spectral technique is to extract useful information from the network traffic signal and to identify suspicious behaviour on the network traffic [3, 4]. A rich scientific contribution exists in this research field. However, challenges are always raised to overcome the existing limits. Recently, and with the evolution of numerical modeling techniques, IDS have been more emphasized on artificial intelligence techniques. In [4], a machine learning method was used to detect Distributed Denial of Service (DDoS) attacks; the proposed method compared to SVM technique demonstrated good detection accuracy and training time. As we know, it has been shown that wavelets are very strong tools for signal processing and are better than Fourier transforms. Wavelet theory allows a signal to be decomposed in several signals representing different frequencies. In recent years, wavelet transforms were used as techniques to build anomaly-based intrusion detection systems to increase the detection rate and to decrease the false alarm. In this chapter, we explore some recent IDS works using wavelet transforms coupled to neural networks and machine learning aiming to combine the advantages of both techniques. The rest of the chapter is organized as follows. Section 2 deals with a brief introduction to multiresolution analysis and wavelet concept. Section 3 presents an overview on neural networks. The next section is devoted to intrusion detection systems based on wavelet transforms. We conclude with a conclusion and some ideas for a future work.

2. Multiresolution Analysis and Wavelets

Beginning from the 1980s, wavelets which are mathematical functions have been used in the mathematic field for signal processing, image compression, partial differential equations solving; and in other fields such as electrical engineering and seismic geology ([5, 6]). In the recent years, wavelets coupling to neural networks, have been applied in the cybersecurity field for intrusion detections systems in order to protect networks from intrusions and from cyber-attacks. Classical approaches to wavelet construction deal with multiresolution analysis (MRA). A wavelet basis set starts with two orthogonal functions: the scaling function or father wavelet ϕ and the wavelet function also called the mother wavelet ψ . A well-known and a simplest type of wavelets named the Haar wavelet was applied by the FBI (Federal Bureau of Investigation of the USA) in the extraction of the characteristics of individual thumb impressions. In [8], we define a multiresolution analysis as a sequence of embedded spaces of approximation of $L^2(\mathbb{R})$, $V_j, j \in \mathbb{Z}$ ($\mathbb{R}$ is the set of real numbers and $\mathbb{Z}$ is the set of positive integers) satisfying the following properties. The spaces V_j are generated by the orthonormal bases $\{\phi_{jk}, k \in \mathbb{Z}\}$ where $\phi_{jk}(x) = 2^j \phi(x - k)$ and ϕ is r -regular. A multiresolution analysis is related to particular functions called wavelets obtained by dilation and translation of a given function ψ , the mother wavelet. Wavelets are essentially fast decaying and localized in time and frequency. Let now W_j be the orthogonal complement of V_j in $V_{j+1} : V_{j+1} = V_j \oplus W_j$, it exists a wavelet ψ such that the family $\{\psi_{jk}(x) = 2^j \psi(x - k), k \in \mathbb{Z}\}$ is an orthonormal basis of W_j . The function ψ has the same regularity properties as ϕ and, moreover, the function ψ has zero moments, i.e. For all real α such that $\alpha \leq r$, $\int_{\mathbb{R}} x^\alpha \psi(x) dx = 0$, and for all function $f \in L^2(\mathbb{R})$ (in other words, $\int_{-\infty}^{\infty} |f(x)|^2 dx < +\infty$), one has the following equality:

$$f(x) = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \langle f, \psi_{jk} \rangle \psi_{jk}(x) \quad (2.1)$$

In the other hand, orthogonal or biorthogonal multiresolution analysis lead naturally to hierarchical

Figure 1: A Meyer Wavelet and its corresponding scaling function: (image source <https://www.mathworks.com/>)

algorithms for the expression of the scalars $\langle u, \psi_{jk} \rangle$ in term of the scalars $\langle u, \phi_{pk} \rangle$, $p > j$, $k \in \mathbb{Z}$. For practical implementations, it implies the existence of filters H_j and G_j that are used in the following way:

$$\langle u, \psi_{jk} \rangle = \sum_{l \in \mathbb{Z}} \langle u, \phi_{j+1,l} \rangle G_j(2k-l), \langle u, \phi_{jk} \rangle = \sum_{l \in \mathbb{Z}} \langle u, \phi_{j+1,l} \rangle H_j(2k-l) \quad (2.2)$$

In the numerical applications, these relations are implemented using convolution and decimation algorithms of complexity $O(N)$ or $O(N \log(N))$, where N is the number of variables. The operator $\{\langle u, \phi_{pk} \rangle, k \in \mathbb{Z}\} \mapsto \{\langle u, \psi_{jk} \rangle, j < p, k \in \mathbb{Z}\}$ is called the wavelet decomposition, while the inverse operator is called the wavelet reconstruction.

To close this section, we remind that there exists a multitude of wavelet types such as the wavelets of Morlet, Meyer, Daubechies, and so one. Figure 1 illustrates an example of Meyer wavelets. The problem that will arise in the construction of an IDS using the wavelet transform will be the choice of the appropriate wavelet; this will require the test of several types of orthogonal and bi-orthogonal wavelets with respect to the specificity of the network traffic signal.

3. Definition of neural networks

Neural networks, also called artificial neural networks, are simple imitations of a neuron in the human brain, devoted to solve machine learning problem where the neuron is considered as a unit which is generally expressed by an activation function (sigmoid function, etc.). Neural network architectures can be divided into four main families: Feed Forwarded Neural Networks, Recurrent Neural Networks (RNN), Resonance Neural Networks and Self-organized Neural Networks. During the last decades, neural networks have been deployed heavily for deep learning techniques to classify objects or to make predictions. In deep learning, data are changing through layers of interconnected nodes where each node is a perceptron acting with a linear regression and a given activation function.

The process of neural networks can be summarized into two major steps: The first one consists on applying a linear transformation. For a given layer and n features, the inputs $x_i, i = 1, \dots, n$ undergo a linear transformation through given weights w_i and a bias b . In other words, the transformation will be in the following form: $x = \sum_{i=1}^n x_i w_i + b_i$. The second step is the call of the activation function that will be applied to the outputs of the first step; The activation function aims to define the output of the neuron in a range of values. There is a variety of the activation functions as for instance Sigmoid, tanh (hyperbolic tangent) or ReLU (Rectified Linear Unit).

To close this paragraph, we can say that various applications of neural networks can be found in image recognition, classifications of texts or images, identification of objects, data prediction, and filtering data sets. The main challenges are to reduce the training time and to achieve highest accuracy. An illustration of a simple neural network is given in figure 2 with input, hidden and output layers.

Figure 2: Illustration of a simple neural network.

Figure 3: The Sigmoid activation function.

3.1. A slight reminder on activation functions

The Sigmoid is a function that makes the value of the outcome of a neuron between 0 to 1 as illustrated on figure 3.1; the mathematical expression of the sigmoid function corresponds to

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (3.1)$$

The tanh activation function plotted on figure 4 corresponds to:

$$\frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (3.2)$$

And finally, the ReLU is a piecewise linear function defined as the positive part of its argument as illustrated on figure 3.3; it has a simple structure. It may led to excellent performance If $g(x)$ defines the function ReLU, then

Figure 4: The Tanh activation function.

Figure 5: The ReLU activation function.

$$g(x) = \max(0, x) \quad (3.3)$$

We remark that both Sigmoid and tanh activation functions suffer from a serious problem of vanishing gradient; from the literature, we can say that (in some applications) tanh could be better than the sigmoid function. In addition, ReLU does not suffer from the vanishing gradient problem. For signal processing applications, the hyperbolic tangent tanh and the logistic sigmoid functions are the most common choices [15].

4. Wavelet based intrusion detection systems

4.1. A brief description of the intrusion detection systems

An intrusion detection system is the process of monitoring systems and networks for abnormal activities or intrusion attempts. It collects events and reports them to the administrator. Intrusion detection systems are mechanisms that enhance the security when implemented in parallel with anti malware, firewalls and access controls. The IDS can be software or hardware aiming to automate the process of monitoring events and to reveal intrusions. They can be categorized into different groups depending on the type of the events they monitor or on how they are deployed: Network-based IDS, Host-based IDS, Hybrid-based IDS, Signature-based IDS, Anomaly-based IDS, Passive-based IDS, and Active-based IDS [4, 9, 10]. An illustration of an IDS on a network architecture is given on figure 6.

4.2. A review on wavelet based intrusion detection systems

This section presents a list of the most relevant works on intrusion detection systems based on wavelets and neural network techniques. The main published works have successfully demonstrated the importance of combining the advantages of neural network and wavelets to build efficient IDS.

In [11], authors presented an IDS designed by the coupling of discrete wavelet transforms and a three layer Artificial Neural Network (ANN). The ANN method was used for classification, and the experimental simulations were performed on KDD99 dataset demonstrating a higher accuracy. In this study, the wavelets were used for data pre-processing denoising information and helping in reducing false alarms. The proposed IDS was recommended for real-time networks.

The work in [12] presented a method of wavelet-analysis combined to the Bayesian method for network traffic time-series for detecting attacks on the digital manufacture infrastructure. The approach gave promising results on detecting DOS attacks like the SlowLoris attacks which consist in opening connections to the target web server. Also, the approach demonstrated a good efficiency in detecting HTTP DOS attacks which overwhelm a targeted server with the HTTP requests.

In [13], an IDS was designed by a combination of the wavelet transform and the artificial neural network which was trained on the KDD 9 dataset. The experimental phase focused on an ad hoc wireless network where five TCP / RPC attacks were carried out. The processing of the received packet traffic was

Figure 6: A network architecture secured with an IDS (Intrusion Detection System).

performed using the Daubechies wavelets. The results obtained a high detection rate, and confirmed well the feasibility of the approach.

The work presented in [14] deals with an IDS defined with the machine learning technique known as SVM (Support Vector Machine) method, coupled to a discrete wavelet transform. The wavelet transform was applied to training data for a feature extraction. These data were the input data for the machine SVM. Otherwise, SOFTMAX was trained to classify the input data into two labels. The experiments were implemented using MATLAB on a dataset of 175,341 instance. Finally, the experimental tests demonstrated satisfactory results when compared with existing results in the field.

5. Discussion and Conclusion

Neural networks based IDS's have been widely studied and deployed in order to prevent cyberattacks. Some of them exhibit good detection rates and low false alarms. DARPA 1999 was the most used dataset for performing experimental results. There is still a challenge for improving the processing time for the intrusion detection, to decrease the training time and to increase the accuracy tests. On the other hand, it is important to point out the need of improving the efficiency of the neural network algorithms, and more particularly the efficiency of the classification methods. Finally, the wavelet transform methods used as pre-processing techniques remain an interesting tool to be explored for the good analysis of the Internet traffic, denoising, extracting the useful information and reducing the false alarms. Thanks to the promising results encountered in the literature, it would be more interesting to deepen the investigation on how the choice of the used wavelets can affect the training time and the false alarms.

References

- [1] Roberto Di Pietro and Luigi V. Mancini. Intrusion Detection Systems. I SBN-13: 978-0387772653, ISBN-10: 9780387772653. Springer 2008 [1](https://doi.org/10.1016/j.comnet.2008.10.001)
<https://doi.org/10.1016/j.comnet.2008.10.001>
- [2] Brij B. Gupta and Srivathsan Srinivasagopalan. Intrusion Detection Systems. ISBN13: 9781799822424, DOI: 10.4018/978-1-7998-2242-4 (ebook) 2020 [1](#)
- [3] X. He, C. Papadopoulos, J. Heidemann, Urbashi Mitra and Usman Riaz. Remote Detection of Bottleneck Links Using Spectral and Statistical Methods. Computer Networks, 53 (3), pp. 279-298, February, 2009 [1](#)
- [4] Chi Cheng, Wee Peng Tay and G. Huang, Extreme learning machines for intrusion detection, The 2012 International Joint Conference on Neural Networks (IJCNN), Brisbane, QLD, 2012, pp. 1-8, DOI: 10.1109/IJCNN.2012.6252449 <https://doi.org/10.1109/ijcnn.2012.6252449>
- [5] Saiida Lazaar et al. Wavelet algorithms for numerical resolution of partial differential equations. Computer methods in applied mechanics and engineering 116 (1-4), 309-314, 1994 [1](#), [4.1](#)
- [6] Stéphane Mallat. A Wavelet Tour of Signal Processing. ISBN 978-0-12-374370-1. Elsevier 2009 [2](#)
- [7] Ingrid Daubechies. Ten Lectures on Wavelets (CBMS-NSF Regional Conference Series in Applied Mathematics). ISBN-13: 978-0898712742, ISBN-10: 0898712742, SIAM 1992 [2](#)

- [8] Y. Meyer. Wavelets and operators. Cambridge University Press. Online publication 2009. ISBN 9780511623820 DOI: <https://doi.org/10.1017/CBO9780511623820> <https://doi.org/10.1017/CBO9780511623820>
- [9] Jacob, N.M., & Wanjala, M.Y. (2018). A Review of Intrusion Detection Systems. Global Journal of Computer Science and Technology
- [10] Robert Barnard. Intrusion Detection Systems. eBook ISBN: 9780080943657. Butterworth-Heinemann 1988 2
- [11] Hamid, Y., Shah, F.A. & Sugumaran, M. Wavelet neural network model for network intrusion detection system. Int. j. inf. tecnol. 11, 251–263 (2019) <https://doi.org/10.1007/s41870-018-0225> <https://doi.org/10.1007/s41870-018-0225-x>
- [12] Darya Lavrova¹, Pavel Semyanov, Anna Shtyrkina, and Peter Zegzhda, Wavelet-analysis of network traffic time-series for detection of attacks on digital production infrastructure. SHS Web of Conferences 44, 00052 (2018) <https://doi.org/10.1051/shsconf/20184400052> <https://doi.org/10.1051/shsconf/20184400052>
- [13] E. W. Tavares Ferreira, G. Arantes Carrijo, R. de Oliveira and N. Virgilio de Souza Araujo. Intrusion Detection System with Wavelet and Neural Artificial Network Approach for Networks. Computers. IEEE Latin America Transactions, vol. 9, no. 5, pp. 832-837, Sept. 2011, DOI: 10.1109/TLA.2011.6030997 <https://doi.org/10.1109/tla.2011.6030997>
- [14] Sefer Kurnaz, Israa Abed Obaid. Support Vector Machine (SVM) Based on Wavelet Transform (WT) for Intrusion Detection System (IDS). IJCSMC, Vol. 8, Issue. 2, February 2019, page.13 –19 4.1
- [15] Zhangyang Wang, Yun Fu and Thomas S. Huang. Deep Learning Through Sparse and Low-Rank Modeling, ISBN 978-0-12-813659-1, 2019 Elsevier Inc. 4.1