New Intrusion Detection System Based on Support Vector Domain Description with Information Gain Metric

Mohamed El Boujnouni and Mohamed Jedra
(Corresponding author: Mohamed El Boujnouni)
Laboratory of Conception and Systems, Mohammed V University
Avenue Ibn Battouta B.P. 1014, Rabat, Morocco
(Email: med_elbouj@yahoo.fr)
(Received Nov. 20, 2016; revised and accepted Feb. 21 & Mar. 4, 2017)

Abstract

With the vulgarization of Internet, the easy access to its resources and the rapid growth in the number of computers and networks, the security of information systems has become a crucial topic of research and development especially in the field of intrusion detection. Techniques such as machine learning and data mining are widely used in anomaly-detection schemes to decide whether or not a malicious activity is taking place on a network. This paper presents a new intrusion detection system (IDS) based on information gain criterion to select relevant features from network traffic records and a new version of support vector domain description to classify the extracted features and to detect new intrusions. Experimental evaluation on NSL-KDD, a filtered version of the original KDD99 has shown that the proposed IDS can achieve good performance in terms of intrusions detection and recognition.

Keywords: Information Gain Metric; Network Intrusion Detection System; NSL-KDD; Support Vector Domain Description

1 Introduction

In computer science domain, an intrusion can be defined as the attempts to compromise the confidentiality, integrity, or availability of a computer or network. Thus, Network Intrusion Detection Systems (NIDS) are a critical defense layer of any network security architecture. The main task of NIDS is to monitor network traffic for suspicious contents, and to alert system administrators when a malicious activity is taken place. The detection of intrusions can be performed basing on analyzing the events which occur in the monitored network. Two primary approaches are used: misuse or signature detection and anomaly detection. The first technique, existing in the majority of commercial NIDSs, aims to detect known attacks by using predefined attack patterns and signatures so it looks for a specific event that has already been recognized and registered. The second technique detects attacks by comparing the deviation from a model describing the normal behavior of the monitored resource. There are advantages and disadvantages associated with each approach: Misuse detection methods can detect malicious network traffics without generating high false alarms but they are basically limited to known attacks. This leads to the necessity for frequent updates of the intrusions database. On the contrast anomaly detection methods based on heuristics or rules are able to detect known and unknown attacks. This propriety is very important since new kinds of vulnerabilities and intrusions are constantly appearing. However, new legitimate behavior can be falsely identified as malicious, resulting in a false positive. Recently new hybrid intrusion detection systems that exploit benefits of both misuse and anomaly detection techniques are developed and showed great success [21, 28].

Anomaly detection approach is based on techniques such as: Threshold detection, rule-based measures, statistical measures, machine learning and data mining methods. The first technique expresses some attributes of user and system behavior in terms of counts. Then it compares the latter with a tolerance level. The second approach tries to define a set of rules that can be used to decide whether a given behavior is normal or not. Statistical measures analyze the distribution of the network traffic attributes and can be parametric or non-parametric, the first one is assumed to fit a particular pattern while the second is learned from a set of historical values. The last technique based on machine learning and data mining learns from a set of training data and constructs a model able to classify new network traffic as legitimate or malicious.

In this paper we aim to design a new intrusion detec-

tion system based on the last technique described above. The proposed NIDS works in three steps: At first, a data encoding and normalization operations are performed on the network traffic records. Then, information gain (IG) method is applied to extract relevant features from the preprocessed data. Finally, a new version of SVDD called SVDD with small sphere and parametric volume (SSPV-SVDD) [3] will be trained with the extracted features and used as a novelty detection model able to detect unknown attacks. Experimental evaluation of our approach will be performed using NSL-KDD a benchmark dataset widely used to evaluate the performance of NIDSs.

This paper is organized as follows: Section 2 presents an overview of some previous applications of machine learning and soft computing methods to detect network intrusions. Section 3 describes in details our new network intrusion detection system. This section is divided into three parts: The first one describes the architecture of the proposed NIDS, the second presents the techniques of data encoding, data normalization and relevant feature extraction and finally the third presents the application of SSPV-SVDD to detect network intrusions. The last section investigates empirically the performance of the proposed NIDS using NSL-KDD. This section is divided into two parts: The first one describes NSL-KDD dataset and the second presents the experimental setting and the results of applying the proposed NIDS on NSL-KDD. A conclusion is provided in the final section.

2 Related Work

There are numerous important research papers regarding the use of machine learning and soft computing techniques to detect network intrusion. For example Liu et al. [18] proposed a genetic clustering method for intrusion detection. Their method is able to establish clusters automatically and to detect attacks by labeling normal and abnormal groups. Javadzadeh and Azmi [13] proposed a hybrid approach to design NIDSs. Their method is able to generate fuzzy rules based on a fuzzy genetic machine learning algorithm and to detect multiple attacks. Aghdam and Kabiri [1] focused on feature selection of network traffic for intrusion detection purpose. They proposed a new method based on ant colony optimization (ACO) algorithm and nearest neighbor classifier to eliminate irrelevant and redundant features from network traffic records. Wang et al. [29] presented an application of artificial neural networks (ANN) and fuzzy clustering on intrusion detection. Their approach works sequentially: Firstly fuzzy clustering technique was applied to create different training subsets. Then, based on the latter, different ANN models are trained to formulate different base models. Finally, a fuzzy aggregation module is employed to aggregate these results. Li et al. [35] introduced an application of multiple kernel support vector machine (SVM) for intrusion detection. This new version of SVM improves the standard one by calculating the weights of kernel functions and Lagrange multipliers simultaneously and automatically without user intervention. Mukherjee and Sharma [20] presented an intrusion detection method based on naive Bayes classifier with a new feature reduction method. In order to select the most relevant features the authors investigate the performance of three standard features selection methods, namely correlation-based features selection, information gain and gain ratio. Then they proposed a new features reduction method named feature vitality. The reduced data sets are further classified using Native Bayes classifier. Li et al. [17] proposed the use of K-means clustering and particle swarm optimization (PSO) algorithm to deal with network intrusions. The key idea behind using PSO is to reach a good overall convergence and to overcome falling into local minima. Wankhade et al. [30] discussed the development of a secured information system by applying various data mining techniques on intrusion detection systems for the effective identification of both known and unknown attacks. Tao et al. [24] presented one-class classification approach to detect network intrusions based on SVDD. They used genetic algorithm to determine the optimal parameter of the kernel function. Then they analyzed the behavior of the classifier basing on the selected parameters. Yu Zhang et al. [36] proposed an optimized method of SVDD based on particle swarm optimization algorithm (PSO). Their method adopts PSO to eliminate the superfluous parameters in SVDD and carries out dimension reduction to data. GhasemiGol et al. [9]. presented a novel approach to describe the normal behavior of computer networks using minimal hyper-ellipse instead of hypersphere used by SVDD. The hyper-ellipse creates tighter boundary around the positive examples. The boundary was used to detect new attacks. Zhou et al. [38] presented an improved intrusion detection method based on kernel learning. They used Kernel principal component analysis (KPCA) as preprocessor of the dataset. Then they applied SVDD on the preprocessed data. Kenaza et al. [16] introduced an adaptive SVDD-based learning for false alarm reduction in intrusion detection. In their work they aimed to take into consideration the dynamic aspect of a monitored environment, and they proposed an adaptive SVDD-based learning approach that aims at continuously enhancing the performances of the SVDD classifier by refining the training dataset. Yang et al. [32] proposed a new method for anomaly intrusion detection based on SVDD. In their work they considered intrusion detection problem as one-class classification and then they built SVDD model for normal data. This model was used to detect known and unknown attacks. Yang et al. [33] introduced a new framework for adaptive anomaly detection based on SVDD classifier and change detection algorithm. The proposed framework consists of four main components: preprocessor, change detector, model generator and anomaly detector.

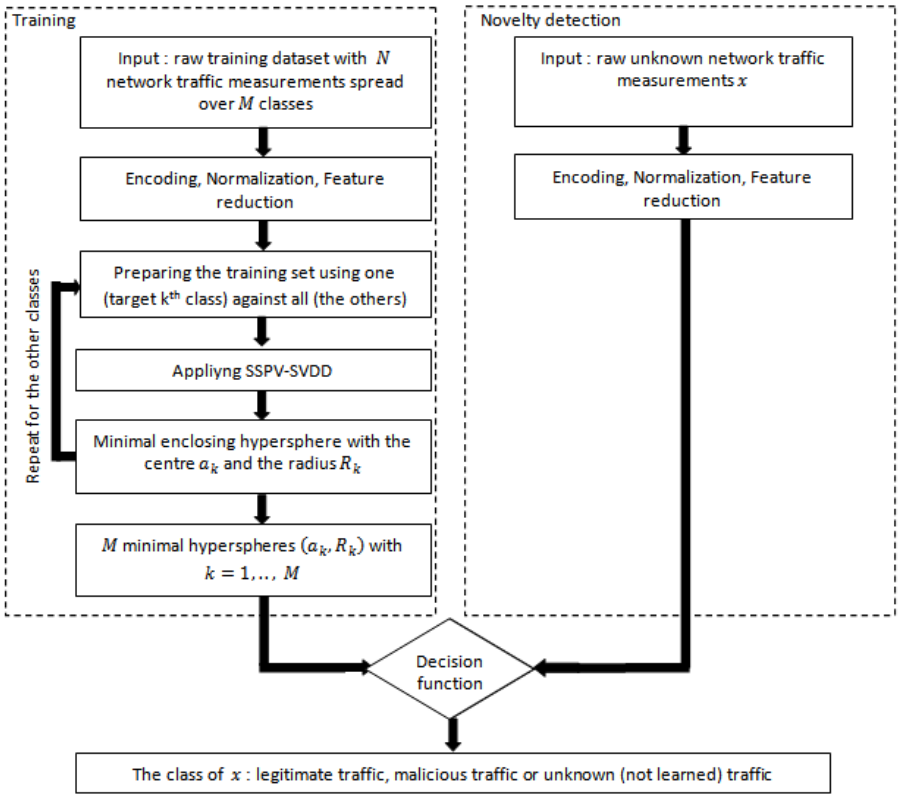

Figure 1: Architecture of the proposed network intrusion detection system

3 The Proposed Network Intrusion Detection System

3.1 Architecture of the Proposed NIDS

Figure 1 shows the architecture of the proposed network intrusion detection system. The proposed NIDS works in two steps:

- Training: After preprocessing (encoding, normalization and feature reduction) a set of network traffic measurements having M classes describing normal and attacks behaviors, SSPV-SVDD with Gaussian kernel will be applied on each class. The result is a set of M minimal hyperspheres each of which has a center a_k and a radius R_k with $k = 1, \cdots, M$ and encloses the samples of a specified class.

- Novelty detection: After preprocessing an unknown network traffic measurement x with the same technique used in the training step. The decision function expressed by Equation (5) will be evaluated. The result is either the class label of x or no one of the learned classes which signify that x is a new type of attacks.

3.2 Encoding, Normalization, and Feature Reduction

Network traffic contains different forms of data (continuous, discrete and symbolic) with significantly varying resolution and ranges, in order to handel this dataset with SSPV-SVDD a preprocessing is required. The latter is based on 3 steps:

Step 1: Convert symbolic attributes to numeric values. The conversion is performed using the encoding tables shown in Tables 1, 2, 3, 4.

Table 1: Encoding of symbols in the 2^{nd} field of NSL-KDD dataset

Symbol	tcp	udp	icmp
Code	1	2	3

Step 2: Normalize numeric values [37]. The data attributes are scaled to fall within the interval $[x_{min}^{new}, x_{max}^{new}]$ that can be [-1, 1] or [0,1].The scaling is performed using Equation (1). Likewise, before testing, the same way is applied to scale testing data. The main advantage is to avoid attribute in greater numeric ranges dominate those in smaller numeric

Table 2: Encoding of symbols in the 3^{rd} field of NSL-KDD dataset

Symbol	ftp_data	other	private	http	remote_job	name	netbios_ns	eco_i
code	1	2	3	4	5	6	7	8
Symbol	mtp	telnet	finger	domain_u	supdup	uucp_path	Z39_50	smtp
code	9	10	11	12	13	14	15	16
Symbol	csnet_ns	uucp	netbios_dgm	urp_i	auth	domain	ftp	bgp
code	17	18	19	20	21	22	23	24
Symbol	ldap	ecr_i	gopher	vmnet	systat	http_443	efs	whois
code	25	26	27	28	29	30	31	32
Symbol	imap4	iso_tsap	echo	klogin	link	sunrpc	login	kshell
code	33	34	35	36	37	38	39	40
Symbol	sql_net	time	hostnames	exec	ntp_u	discard	nntp	courier
code	41	42	43	44	45	46	47	48
Symbol	ctf	ssh	daytime	shell	netstat	pop_3	nnsp	IRC
code	49	50	51	52	53	54	55	56
Symbol	pop_2	printer	tim_i	pm_dump	red_i	netbios_ssn	rje	X11
code	57	58	59	60	61	62	63	64
Symbol	urh_i	http_8001	aol	http_2784	tftp_u	harvest		
code	65	66	67	68	69	70		

Table 3: Encoding of symbols in the 4^{rd} field of NSL-KDD dataset

Symbols	SF	S0	REJ	RSTR	SH	RSTO	S1	RSTOS0	S3	S2	OTH
Codes	1	2	3	4	5	6	7	8	9	10	11

Table 4: Encoding of symbols in the 41^{rd} field of NSL-KDD dataset

Class	**Sampling Rate**	**Length**	**Code**
Non-attack	1	Normal	0
DOS	10	Back, land, neptune, pod, smurf, teardrop, apache2, processtable, worm, udpstorm, mailbomb	1
Probe	6	Ipsweep, portsweep, nmap, satan, saint, mscan	2
R2L	16	Warezclient, guess_passwd, ftp_write, multihop, imap, warezmaster, phf, spy, snmpgetattack, httptunnel, snmpguess, named, sendmail, xlock, xsnoop	3
U2R	7	Rootkit, buffer_overflow, loadmodule, perl, ps, xterm, sqlattack	4

ranges.

$$x^{new} = x^{new}_{min} + \frac{x^{new}_{max} - x^{new}_{min}}{x^{old}_{max} - x^{old}_{min}}(x^{old} - x^{old}_{min}). \quad (1)$$

With x^{old}_{max} and x^{old}_{min} are respectively the maximum and the minimum values of the attribute that x belongs to, x^{old} is the value before normalization and x^{new} is the value after normalization that will belong to the interval $[x^{new}_{min}, x^{new}_{max}]$.

Step 3: Extract relevant features using information gain measure that can be expressed as follows: Let S be a set of M classes that contains s labeled training points where each class I includes s_i samples. Expected information needed to classify a given sample is evaluated using the following equation:

$$I(s_1, s_2, ..., s_M) = -\sum_{i=1}^{M} \frac{s_i}{s} log_2(\frac{s_i}{s}).$$

An attribute A with values $\{A_1, A_2, .., A_v\}$ can split the training set S into v subsets $\{S_1, S_2, .., S_v\}$ where S_j is the subset which has the value A_j for attribute A and contains s_{ij} points of class i. The entropy of the attribute A can be expressed as:

$$E(A) = \sum_{j=1}^{v} \frac{s_{1j} + ... + s_{Mj}}{s} I(s_{1j}, s_{2j}, ..., s_{Mj}). \quad (2)$$

Information gain for A is given by the equation:

$$Gain(A) = I(s_1, s_2, , s_M) - E(A). \quad (3)$$

3.3 Application of SSPV-SVDD to Detect Network Intrusion

Support Vector Domain Description is a relatively new classification method inspired by Support Vector Machine (SVMs). SVDD was originally developed by Tax and Duin [26, 27] and then improved by many researchers. This classifier aims to enclose the data of interest through the smallest hypersphere where its boundary serves to classify new unknown samples. Due to its high generalization capability, SVDD have been applied successfully to a wide range of problems, such as: Biometric authentication [10], novelty detection [7, 31], fault diagnosis [6, 34], credit ratings [8, 22], disease diagnosis [5, 15], digital investigations and computer security [4, 19], financial fraud detection [2, 14], etc. SVDD inherits many of the advantages of SVMs, including SVDD has a solid mathematical foundation based on the statistical learning theory. Also, it benefits from kernel functions that maps a linearly inseparable data points represented in the original space into a high dimensional feature space in which they become separable. In addition, training a given dataset with SVDD implies solving a constrained quadratic problem (QP) with a single minimum which avoids the risk of becoming trapped by local minimum solutions. Moreover, the classification of a new unknown sample requires checking the sign of a decision function basing only on a small subset of the training data known as support vectors (SVs) which reduces the time required to classify new unknown instances. Furthermore, training SVDD requires setting a small number of parameters which limits the intervention of users.

The proposed NIDS is designed with an improved version of SVDD called SSPV-SVDD [3]. The latter aims to improve SVDD by introducing a new regularization parameter that offers the following advantages: 1) It allows user to customize the hyperspherical boundary between different classes; 2) It plays a compromise between the acceptance of negative data and the rejection of target data; 3) It allows to distinguish between the set of samples existing on the boundaries.

SSPV-SVDD considers a dataset $S = \{(x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N)\}$ with $i = 1, \cdots, N$ and $x_i \in \mathbb{R}^d$. The label y_i equals $+1$ for the target samples, and -1 for the negative ones. The objective of SSPV-SVDD is to find the smallest hypersphere, with a center a and a radius R that includes the maximum number of target samples and excludes the majority of negative ones following the value of a regularization parameter called p. This problem is formulated as follows:

Minimize:

$$R^2 + C\sum_{i=1}^{N} \varepsilon_i$$

Subject to:

$$\|x_i - a\|^2 \leq R^2 - p.y_i + \varepsilon_i, \forall i = 1, .., N, \quad \text{with } y_i = +1$$

$$\|x_i - a\|^2 \geq R^2 - p.y_i - \varepsilon_i, \forall i = 1, .., N, \quad \text{with } y_i = -1.$$

Where $\|.\|$ is the Euclidean norm. p is a strictly positive real number. ε_i are slack variables that measure the violation amount of the constraints. To allow the presence of outliers a positive parameter C was introduced, the latter gives the tradeoff between the volume of the sphere and the rejection of target samples.

It's an optimization problem with constraints that may be solved by Lagrange's method. The primal problem of SSPV-SVDD can be written as follows:

$$L(R, \varepsilon, a) = R^2 + C\sum_{i=1}^{N} \varepsilon_i - \sum_{i=1}^{N} \varepsilon_i y_i - \sum_{i=1}^{N} \alpha_i y_i (R^2 - \|x_i - a\|^2 - p y_i).$$

Where α_i and μ_i are Lagrange multipliers, Annulling the

partial derivatives of L with respect to R, a, ε_i yields:

$$\frac{\partial L}{\partial R} = 0 \quad \Rightarrow \quad \sum_{i=1}^{N} \alpha_i y_i = 1.$$
$$\frac{\partial L}{\partial a} = 0 \quad \Rightarrow \quad a = \sum_{i=1}^{N} \alpha_i x_i y_i$$
$$\frac{\partial L}{\partial \varepsilon_i} = 0 \quad \Rightarrow \quad \alpha_i = C - \mu_i.$$

The dual optimization problem can be written as:

Maximize:

$$L(\alpha) = -\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j x_i x_j + \sum_{i=1}^{N} \alpha_i (y_i x_i x_i + p)$$

Subject to:

$$0 \leq \alpha_i \leq C \; and \; \sum_{i=1}^{N} \alpha_i y_i = 1. \tag{4}$$

In multi-classes case, each class $k = 1, .., M$ is represented with a small sphere (a_k, R_k). To classify a new unknown sample z we have simply to evaluate the decision function given bellow:

$$class \; of \; z = argmax_{k=1,...,M}(1 - \sqrt{\frac{\|z - a_k\|^2}{R_k^2}})$$
$$with \; \|z - a_k\| \leq R_k.$$

The radius R_k that corresponds to the k^{th} class (minimal sphere), can be expressed as follows:

$$R_k^2 = \|x_s - a_k\|^2 + y_s p. \tag{5}$$

Where x_s is a training point that belongs to the set of Support Vectors of the k^{th} class having $0 < \alpha_s < C$.

4 Experimental Setup and Performance Evaluation

4.1 Description of KDD CUP 99 Dataset

In order to evaluate the performance of our new intrusion detection system, we propose to perform several experiments basing on KDD99 database. The latter is a version of the 1998 DARPA intrusion detection evaluation data set prepared and managed by MIT Lincoln Labs [12]. KDD99, widely used by NIDS researchers, includes a wide variety of intrusions simulated in a network environment. In this dataset, each record is composed by 41 fields and categorized into one of 5 classes that are normal and 4 types of attacks: DoS (denial of service), R2L (root to local), U2R (user to root) and Probing (surveillance).

As the number of records in the original KDD99 is very large and contains redundant data, we propose to work with the new filtered version of the original. This new dataset, called NSL-KDD, was created by Tavallaee et al. [25] after a statistical analysis of KDD99. NSL-KDD has the following major advantages over the original KDD99 [11]. It eliminates redundant records to which the classifier will be biased towards during the training. and it contains a reasonable number of records. Therefore, experiments can be achieved using the whole dataset without the need to randomly select a reduced subset.

Similar to KDD-99, NSL-KDD has the same number of attributes with a reduced number of samples. For KDD99 the training dataset contains 494021 records, and the testing dataset consists of 311029 records while NSL-KDD contains 125973 patterns for training and 22544 patterns for testing.

Figure 2 shows a random record taken from NSL-KDD. It can be seen that each attribute can take one of the following forms: continuous, discrete, or symbolic. The before last attribute in each record describes the class label that can take 40 values normal and 39 attacks that can be classified in 4 main intrusions DoS, R2L, U2R and Probing.

```
0,tcp,ftp_data,SF,491,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,2,2,0.00,
0.00,0.00,0.00,1.00,0.00,0.00,150,25,0.17,0.03,0.17,0.00,0.00,
0.00,0.05,0.00,normal,20
```

Figure 2: Random record taken from NSL-KDD dataset

4.2 Experimental Results

The intrusion dataset NSL-KDD is divided on two subsets training and testing. In the experiment the classifier SSPV-SVDD will be trained by the first subset, and then tested by the second. The parameter C is fixed at 100 and the kernel is Gaussian with $\sigma \in \{0.15, 0.30, ..., 1.5\}$. Gaussian kernel was chosen because of its wide uses in pattern recognition problems and its good generalization capability. Since we are using the whole NSL-KDD dataset having 125973 records the traditional QP solvers can 't be applied directly because they need to store Gram matrix of size N×N (125973×125973). To deal with this problem we propose to solve the QP of SSPV-SVDD with the algorithm called Sequential Minimal Optimization (SMO) [23], SMO is an iterative algorithm that decomposes the Quadratic Problem given by Equation (4) to the extreme in such a way that the working set only has two samples and their optimal Lagrange multipliers can be solved analytically. The recognition rates are calculated for the training and testing sets using the following

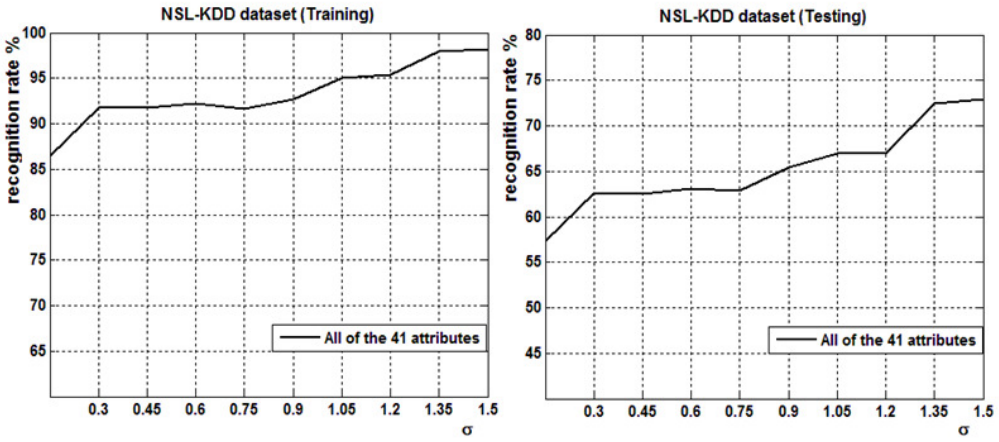

Figure 3: Classification accuracy of NSL-KDD using SSPV-SVDD with 41 attributes

equations:

$$\%(Normal) = \frac{\#of\ normal\ records\ well\ classified}{Total\ number\ of\ records} \times 100$$

$$\%(DoS) = \frac{\#of\ attacks\ DOS\ well\ classified}{Total\ number\ of\ records} \times 100$$

$$\%(R2L) = \frac{\#of\ attacks\ R2L\ well\ classified}{Total\ number\ of\ records} \times 100$$

$$\%(U2R) = \frac{\#of\ attacks\ U2R\ well\ classified}{Total\ number\ of\ records} \times 100$$

$$\%(Probing) = \frac{\#of\ attacks\ Probing\ well\ classified}{Total\ number\ of\ records} \times 100$$

$$Global\ recognition\ rate = \%(Normal) + \%(DoS) + \%(R2L) + \%(U2R) + \%(Probing).$$

Figure 3 shows the classification accuracy of NSL-KDD dataset using the entire 41 attributes. The figure is divided into two parts: The left side shows the global recognition rate of the training dataset with different values of σ. It can be seen that the recognition rate grows with σ until a maximum value that reaches 98%. This signifies that relatively 123454 training instances out of 125973 are well enclosed by the minimal five hyperspheres. The right side illustrates the generalization capability of our NIDS. It can be observed that the recognition rate increases with the kernel width until a maximum of 72.5% which means that 16344 of records out of 22544 are well classified. To increase further the recognition rate next experiment will be performed with the most significant attributes instead of the all ones.

Figure 4 shows the 41 attributes of NSL-KDD dataset sorted in descending order of their information gain measurement, the latter is evaluated using Equations (2) and (3). The main objective of this experiment is to reduce the number of NSL-KDD attributes by selecting the most relevant ones. This will decrease the space and time complexities required to solve the QP of SSPV-SVDD expressed by Equation (4) and could obtain a tight description of NSL-KDD dataset. By analyzing the histogram, we observe that the IG differs from an attribute to another and the last IGs are practically nulls. We will choose to eliminate the attributes with IG<0.001.

Figure 5 shows the classification accuracy of NSL-KDD dataset using the attributes having IG≥0.001. Also, the figure is divided into two parts: The left side describes the recognition rate of the training dataset. It can be seen that the recognition rate is better than the previous experiment and it reaches 99%. The right side represents the novelty detection capability of our NIDS. It can be observed that the recognition rate grows with σ until a maximum value which outperforms the previous experiment and reaches a maximum of 77.5%. This means that the selection of the most significant attributes of NSL-KDD using information gain metric was performed successfully and have improved the classification accuracy of our NIDS.

5 Conclusions

In this paper, we have presented a new network intrusion detection system based on anomaly detection approach. The proposed system includes: Data transformation where symbolic attributes of network traffics are converted to numeric, normalization operation where the numerical attributes are scaled in a small specified range, relevant attributes selection where information gain method was applied as a measure to estimate the quality of the attributes, and finally a novelty detection model based on SSPV-SVDD as classifier and SMO as solver to decide whether a network traffic is an attack or normal.In contrast to numerous IDSs researchers who use just a small random subset of NSL-KDD in the experimental evalu-

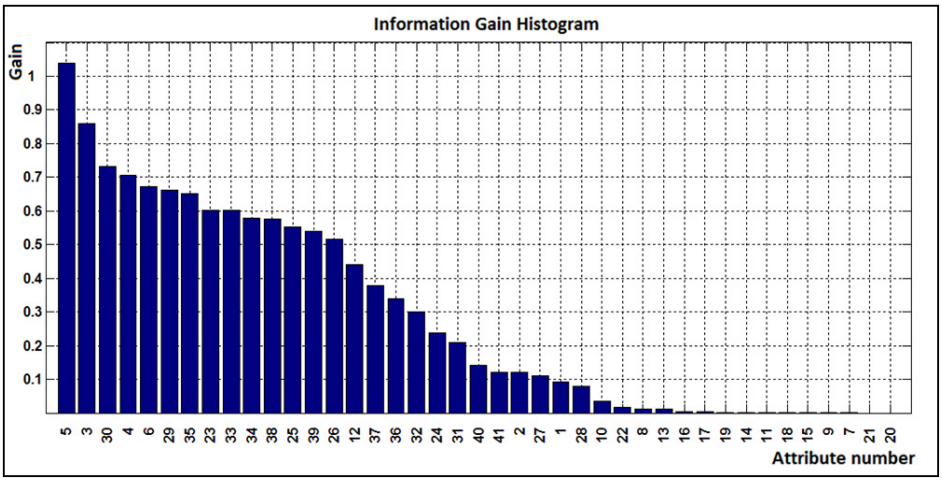

Figure 4: The 41 attributes of NSL-KDD shorted in descending order of IG

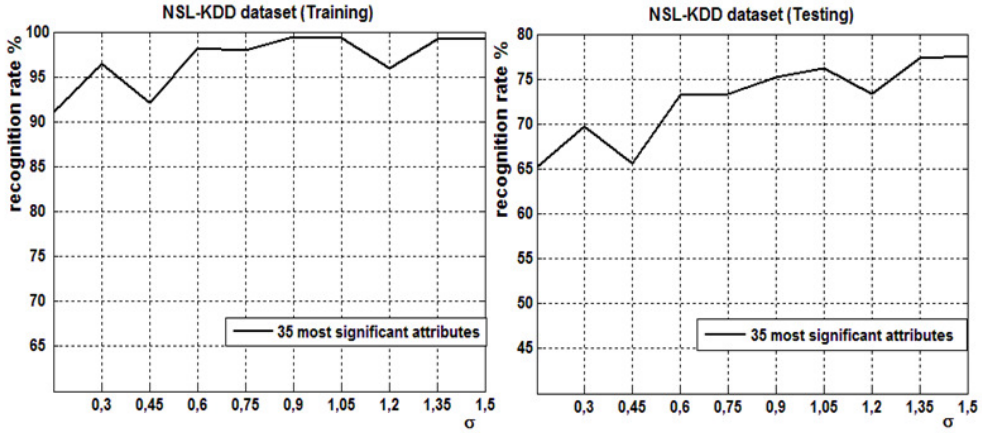

Figure 5: Classification accuracy of NSL-KDD using SSPV-SVDD with the attributes having IG$\geq$ 0.001

ation which gives good but inexact results, in this work we have tested our IDS with the whole NSL-KDD which contains 125973 of samples for training and 22544 samples for testing. The experimental results have shown that with the most significant attributes of NSL-KDD, the proposed IDS can learn 124713 network traffics and can classify successfully 17471 of unknown network behaviors which gives 77.5% of novelty detection rate. This proves that the proposed NIDS is efficient and accurate in detecting different kinds of attacks..

References

[1] M. H. Aghdam and P. Kabiri, "Feature selection for intrusion detection system using ant colony optimization," *International Journal of Network Security*, vol. 18, no. 3, pp. 420–432, 2016.

[2] S. H. An, K. Nam, M. K. Jeong, and Y. R. Choi, "User action-based financial fraud detection method by svdd," *International Journal of Security and Its Applications*, vol. 10, no. 2, pp. 247–254, 2016.

[3] M. El Boujnouni, M. Jedra, and N. Zahid, "A small sphere and parametric volume for support vector domain description," *Journal of Theoretical and Applied Information Technology*, vol. 46, no. 1, pp. 471–478, 2012.

[4] M. El Boujnouni, M. Jedra, and N. Zahid, "New malware detection framework based on n-grams and support vector domain description," in *Proceedings of the* 11^{th} *International Conference on Information Assurance and Security (IAS'15)*, pp. 123–128, Marrakesh, Morocco, Dec 2015.

[5] J. Cao, L. Zhang, B. Wang, F. Li, and J. Yang, "A fast gene selection method for multi-cancer classification using multiple support vector data description," *Journal of Biomedical Informatics*, vol. 53, pp. 381–389, 2015.

[6] L. Duan, M. Xie, T. Bai, and J. Wang, "A new support vector data description method for machinery fault diagnosis with unbalanced datasets," *Expert Systems with Applications*, vol. 64, no. 1, pp. 239–246, 2016.

[7] P. Duong, V. Nguyen, M. Dinh, T. Le, D. Tran, and W. Ma, "Graph-based semi-supervised support vector data description for novelty detection," in *Proceedings of the International Joint Conference on Neural Networks*, pp. 1–6, Killarney, Ireland, July 2015.

[8] C. Gangolf, R. Dochow, G. Schmidt, and T. Tamisier, "Svdd: A proposal for automated credit rating prediction," in *Proceedings of the International Conference on Control, Decision and Information Technologies (CoDIT'14)*, pp. 48 –53, Metz, France, Nov. 2014.

[9] M. GhasemiGol, R. Monsefi, and H. S. Yazdi, "Intrusion detection by ellipsoid boundary," *Journal of Network and Systems Management*, vol. 18, no. 3, pp. 265–282, 2010.

[10] Y. Guerbai, Y. Chibani, and B. Hadjadji, "The effective use of the one-class svm classifier for handwritten signature verification based on writer-independent parameters," *Pattern Recognition*, vol. 48, no. 1, pp. 103–113, 2015.

[11] L. M. Ibrahim, D. T. Basheer, and M. S. Mahmod, "A comparison study for intrusion database (kdd99, nsl-kdd) based on self organization map (som) artificial neural network," *Journal of Engineering Science and Technology*, vol. 8, no. 1, pp. 107–119, 2013.

[12] KDD, *UCI KDD Cup 1999 Data The UCI KDD Archive Information and Computer Science*, University of California Irvine, 1999. (`http://kdd.ics.uci.edu/databases`)

[13] G. Javadzadeh and R. Azmi, "Idufg: Introducing an intrusion detection using hybrid fuzzy genetic approach," *International Journal of Network Security*, vol. 17, no. 6, pp. 754–770, 2015.

[14] M. K. Jeong, S. H. An, and K. Nam, "Svdd-based financial fraud detection method through respective learnings of normal/abnormal behaviors," *International Journal of Security and Its Applications*, vol. 10, no. 3, pp. 429–438, 2016.

[15] R. Chandpa Kalpit and M. Jani Ashwini, "Comparative study between two-class svm and one-class svm classifiers for outlier detection for disease diagnosis," *International Journal of Data Mining And Emerging Technologies*, vol. 5, no. 1, pp. 42–48, 2015.

[16] T. Kenaza, A. Labed, Y. Boulahia, and M. Sebehi, "Adaptive svdd-based learning for false alarm reduction in intrusion detection," in *Proceedings of the* 12^{th} *International Conference on Security and Cryptography*, pp. 405–412, Colmar, Alsace,France, July 2015.

[17] Z. Li, Y. Li, and L. Xu, "Anomaly intrusion detection method based on k-means clustering algorithm with particle swarm optimization," in *Proceeding of the International Conference on Information Technology, Computer Engineering and Management Sciences*, pp. 157–161, Nanjing, Jiangsu, Sept 2011.

[18] Y. Liu, K. Chen, X. Liao, and W. Zhang, "A genetic clustering method for intrusion detection," *Pattern Recognition*, vol. 37, no. 5, pp. 927–924, 2004.

[19] Z. Liu, D. Lin, and F. Guo, "A method for locating digital evidences with outlier detection using support vector machine," *International Journal of Network Security*, vol. 6, no. 3, pp. 301–308, 2008.

[20] S. Mukherjee and N. Sharma, "Intrusion detection using naive bayes classifier with feature reduction," in *Proceeding of the* 2^{nd} *International Conference on Computer, Communication, Control and Information Technology*, pp. 119–128, Hooghly, West Bengal, India, Feb. 2012.

[21] E. U. Opara, O. A. Soluade, "Straddling the next cyber frontier: The empirical analysis on network security, exploits, and vulnerabilities," *International Journal of Electronics and Information Engineering*, vol. 3, no. 1, pp. 10–18, 2015.

[22] S. Pang, S. Li, and J. Xiao, "Application of the algorithm based on the pso and improved svdd for the personal credit rating," *Journal of Financial Engineering*, vol. 1, no. 4, 2014.

[23] J. C. Platt, *Advances in Kernels Methods: Support Vector Learning*. Cambridge, Mass: MIT Press, 1998.

[24] X. Tao, F. Liu, and T. Zhou, "A novel approach to intrusion detection based on support vector data description," in *Proceeding of the* 30^{th} *Annual Conference of IEEE Industrial Electronics Society*, pp. 2016–2021, Busan, South Korea, Nov. 2004.

[25] M. Tavallaee, E. Bagheri, W. Lu, and A. Ghorbani, "A detailed analysis of the kdd cup 99 data set," in *Proceeding of the* 2^{nd} *IEEE Symposium on Computational Intelligence for Security and Defence Applications*, pp. 1–6, Ottawa, Ontario, July 2009.

[26] D. M. J. Tax and R. P. W. Duin, "Support vector domain description," *Pattern Recognition Letters*, vol. 20, no. 11-13, pp. 1191–1199, 1999.

[27] D. M. J. Tax and R. P. W. Duin, "Support vector data description," *Machine Learning*, vol. 54, no. 1, pp. 45–66, 2004.

[28] A. Tayal, N. Mishra and S. Sharma, "Active monitoring & postmortem forensic analysis of network threats: A survey," *International Journal of Electronics and Information Engineering*, vol. 6, no. 1, pp. 49–59, 2017.

[29] G. Wang, J. Hao, J. Ma, and L. Huang, "A new approach to intrusion detection using artificial neural networks and fuzzy clustering," *Expert Systems with Applications*, vol. 37, no. 9, pp. 6225–6232, 2010.

[30] K. Wankhade, S. Patka, and R. Thool, "An overview of intrusion detection based on data mining techniques," in *Proceedings of the International Conference on Communication Systems and Network Technologies*, pp. 626–629, Gwalior, India, Apr. 2013.

[31] H. Wenjun, S. Wang, Y. Liu, F. Chung, and W. Ying, "Privacy preserving and fast decision for novelty detection using support vector data description," *Soft Computing*, vol. 19, no. 5, pp. 1171–1186, 2015.

[32] M. Yang, H. Zhang, J. Fu, and M. Luo, "Anomaly intrusion detection method based on svdd," *Computer Engineering*, vol. 31, no. 3, pp. 39–42, 2005.

[33] M. Yang, H. G. Zhang, J. M. Fu, and F. Yan, *A Framework for Adaptive Anomaly Detection Based on Support Vector Data Description*, LNCS 3222, Springer, 2004.

[34] G. Yin, Y. T. Zhang, Z. N. Li, G. Q. Ren, and H. B. Fan, "Online fault diagnosis method based on incremental support vector data description and extreme learning machine with incremental output structure," *Neurocomputing*, vol. 128, pp. 224–231, 2014.

[35] L. Yuping, L. Weidong, and W. Guoqiang, "An intrusion detection approach using svm and multiple kernel method," *International Journal of Advancements in Computing Technology*, vol. 4, no. 1, pp. 463–469, 2012.

[36] X. Y. Zhang, Z. W. Wei, and X. Lin, "Research on svdd network intrusion detection of the optimal feature selection for particle swarm," *Applied Mechanics and Materials*, vol. 716-717, no. 1, pp. 860–863, 2015.

[37] S. Zheng, H. Tang, Z. Han, and H. Zhang, "Solving large-scale multiclass learning problems via an efficient support vector classifier," *Journal of Systems Engineering and Electronics*, vol. 17, no. 4, pp. 910–915, 2006.

[38] Z. X. Zhou, Y. Jiang, L. T. Ming, M. F. Wang, G. C. Xie, and X. Li, "Improved intrusion detection method based on kernel learning," *Computer Engineering*, vol. 38, no. 14, pp. 21–25, 2012.

Biography

Mohamed El Boujnouni received his Ph.D. degree in Computer Science in July 2015, from Mohammed V University Faculty of Sciences, Rabat, Morocco. His research interests include machine learning, data mining, computational intelligence, and pattern recognition.

Mohamed Jedra holds Doctorat de Troisime Cycle and Doctorat d 'Etat degrees in Electronics Engineering and Informatics; all from Mohammed V University in Rabat, Morocco. From 1990 to 1999, he was the Network and Internet Center Manager in the Faculty of Sciences and Assistant Professor at the Department of Physics. In 1999, he became a Professor Habilit of Informatics in the same Department and in 2003 he was promoted to the position of Professor. From 1987 to the present, he was a member of the Conception and Systems Laboratory. He is the co-founder and the current Chair of the Architecture of Informatic Systems UFR /Formation and Research Unit in the Faculty of Sciences in Rabat since 2003. He is also the co-founder and the current Chair of the Security of Informatic Networks and Embedded Systems Master (ScuRISE) in the same faculty. His main research interests include computational intelligence, pattern recognition and biometric. Mohamed Jedra is a member of IEEE since 1995. He is also a member of IEEE Computer Society, IEEE Computational Intelligence Society and IEEE Signal Processing Society.