Journal of Mobile Multimedia, Vol. 13, No.3&4 (2017) 197-209
© Rinton Press

A NOVEL ANOMALY INTRUSION DETECTION BASED ON SMO OPTIMIZED BY PSO WITH PRE-PROCESSING OF DATA SET

MEHDI MOUKHAFI

Informatics and Applications Laboratory (IA), Department of Mathematics and Computer Science,
Faculty of Sciences, Moulay Ismail University, Meknes, Morocco
mehdi.moukhafi@gmail.com

KHALID EL YASSINI

Informatics and Applications Laboratory (IA), Department of Mathematics and Computer Science,
Faculty of Sciences, Moulay Ismail University, Meknes, Morocco
khalid.elyassini@gmail.com

SEDDIK BRI

Materials and Instrumentations (MIN), Department of Electrical Engineering
Superior School of Technology: ESTM, Moulay Ismail University, Meknes, Morocco
briseddik@gmail.com

Current IDSs are mainly based on techniques based on heuristic rules called signatures to detect intrusions in a network environment. These approaches based signature could only detect a known attacks and referenced above. Since there is no signature for new attacks, other approaches must be taken in consideration, such as algorithms learning machine. However, the major problem of IDSs based on learning machine is the high rate of false positives. This study proposes a novel method of intrusion detection based on pre-processing of training data and a combination PSO (Particle Swarm Optimization) -SMO (Sequential minimal optimization) to develop a model for intrusion detection system. The simulation results show a significant improvement in performances, all tests were realized with the kdd99 data set. compared with other methods based on the same dataset, the proposed model shows high detection performances.

Key words: Computer & Network Security, Intrusion Detection System (IDS), Anomaly Based Intrusion Detection, PSO, SMO, KDD Cup 1999 Dataset.

1 Introduction

The development of networks, the data transfer rate and an unpredictable internet use create more problems and anomalies. This greatly complicates the task of mechanisms security such as IDS (Intrusion Detection System). So, researchers need to develop more reliable and autonomous systems that can carry out the anomaly detection without the need of human interaction.

An intrusion detection system (IDS) was presented for the first time by Anderson in 1980 [1], and later formalized by Denning [2], it can be used in the global security politics, which includes other protection tools, such as firewalls and anti-virus software. Thus, it is important to take the advantage of these tools collaboration and their complementarity**.** The intrusion detection system could exist lonely; in this case it is very important to optimize its exploitation. Nevertheless, in all cases, it is essential to improve the IDS performances. The stability and accuracy of detection are two key indicators used to evaluate IDS [3]. Most of IDS research studies have been realized to improve the stability and accuracy of the detection [4,5].

Actual IDS's based on heuristic rules, named signatures. However, they cannot adequately treat a new types of attacks where constantly the environments were changing, the major drawback of approaches based signature is that they only detect known attacks, which implies a frequent updating of the rules database and the time for the implement.

To solve the above mentioned problems, we present a novel intrusion detection approach combining SMO and PSO to enhance the detection precision. In the proposed method, SMO classifier is employed to estimate whether the action is an attack. In order to improve the performance of SMO classification model, PSO is used to optimize SMO.

This paper is organized as follows: Section 1 related work, section 2 describes the framework background, Section 3 introduction to SMO and PSO methods, Section 4 description of the data set used, Section 5 proposed architecture, section 6 simulation of results and evaluation of the proposed method and Comparison with other methods. Section 7 conclusion and future work.

2 Related work

Current research chose using of knowledge discovery in databases (KDD), allowing by IDS's, to learn automatically the networks behaviors, by analyzing the data tracks and their activities. Learning algorithms can play an important role in detecting attacks (known or unknown). Additionally, the IDS's performances are considerably improved at the network level.

The data mining techniques are better applied equally for anomaly intrusions detection, also for knowledge-based intrusions detection [6]. The statistical analysis of the normal system behavior is one of the first approaches to intrusion detection. The statistics are used mathematically to describe an observed mechanism. Thottan and al. [7] propose a statistical treatment technique of the signal based on the exponential change to detect anomalies in traffic network.

Song and Ling [8] has developed a technique based on aggregate flows that significantly reduces the amount of monitoring data and manages high amounts of statistics and packet data. This technique use the flow measurement mechanism Netflow [9], they collect the data stream and select the five key needed to identify a malicious traffic (Source IP address, destination IP address, port number, source, number of destination port, protocol layer three).

Learning algorithms can play an important role in detecting attacks (known or unknown). Additionally, the IDS's performances are considerably improved at the network level. SVM obtains a good detection performance in terms of classifying the flow of a network into normal or abnormal behaviors. Feng et al. [10] introduced an approach combining SVM with self-organized ant colony network.

Kuang et al. [11] propose a solution based on a combination of the SVM model with kernel principal component analysis (KPCA) and genetic algorithm. KPCA was used to reduce the dimensions of feature vectors, whereas GA was employed to optimize the SVM parameters.

Wathiq et al. [12] propose a solution based on hybrid SVM and Extreme Learning Machine model Learned with data set built by a modified K-means. The modified K-means is used to build new small training datasets representing the entire original training dataset.

Intrusion detection technique used by Saied et al. [13] to detect DDOS attacks known and unknown in real time, based on the use of an artificial neural network (ANN) and specific characteristics (models) that differs between legitimate traffic and traffic attack using learning by back propagation coupled with a sigmoid activation function, the authors has selected three ANN topological structures, one for the most used protocols (TCP, UDP, ICMP) in DDOS attacks, each one with three layers (input, hidden and output). The number of nodes in each topological structure is different, the ICMP topological structure consists of three inputs and four hidden nodes, the topology structure TCP consists of five inputs and four hidden nodes and the topological structure of UDP consists of four inputs and three hidden nodes that treat the calculation process with respect to input and output nodes. The output layer consists output node for the attacks and an output node for legitimate traffic.

Monowar and al. [14] introduce a procedure based on a mutual and general information, an entropy function selection technique for selecting a non-redundant subset of features, based on a clustering and trees to generate a set of reference points and a function of aberrant score to classify incoming network traffic to identify anomalies.

The detection of outlier's technique based on the distance is presented by Knorr et al. [15], it defines a point to a remote outlier if at least a fraction user-defined point in the dataset is further than a certain minimum distance.

Nadiammai and Hemalatha [16] Have established a DOS attack detection with a hybrid mechanism which is based on two steps, the first is the submission of the traffic flow which has an intrusion detection system based signature (SNORT) for preliminary detection of the incoming flow and detecting known attacks, the second step is the recovery of the classified flows as legitimate and applies a classification based supervised learning (SVM) with the use of the kernel function Radial basis function.

The present study proposes an intrusion detection method based on the combination of two algorithms PSO and SMO, and then examine the presented method on the data set KDD 1999, to show that the system performance intrusion detection is significantly promoted by our method. Compared with other methods tested on the same dataset, the proposed model shows high detection rate and its accuracy.

3 THE USED METHODS

3.1 Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a stochastic optimization method, for the nonlinear functions, inspired by the social behaviour of insect colonies, bird flocks, fish schools and other animal society, PSO was invented by Russell Eberhart and James Kennedy [17] in 1995. Originally, the two began developing software simulations birds flocking around food sources, later after realizing that their algorithm solve optimization problems, they present [18] a discrete binary PSO algorithm developed from the previous PSO and operating in continuous variables.

PSO is an iterative algorithm to find the best solution based on a population composed of many particles. For example, a flock of birds (particles) encircling an area where they can feel a hidden source of food. Whoever the closest to food warn others birds to move toward its direction. If any of the other birds circling closer to the target more than the first, it warbles stronger and the others move towards him. This scheme continues until one of the birds find food.

A particle (candidate solution) that may move to the optimal position by updating its position and its speed. The speed of movement of a particle can be updated by the weight of inertia, cognitive learning factor, and the values of social learning factors.

Each particle is an overall potential of the optimum function f (x) on a given area D is considered a point in D-dimensional space and represented as Xi = [xi1, Xi2, . . ., XID] and the velocity vector of the particle is V th = [vi1, vi2, . . ., VD]. In additionally, the best previous position will be replaced by a best fitness value for the particle is PBi th = [pbi1, PbI2, ..., pBID] and the best position to date in the area is GBI = [gb1, gb2, . . ., GBD]. The speed and position of the ith particle is updated according to the equations. (1) and (2):

$$v_{id}^{k+1} = \omega \times v_{id}^{k} + c_1 r_1 \left(pBest_{id}\text{-}x_{id}^{k}\right) + c_2 r_2 \left(gBest - x_{id}^{k}\right) \qquad (1)$$

$$x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1}, \quad d = 1,2,\dots,D; \quad id = 1,2,\dots,N \qquad (2)$$

where v_{id}^{k+1} is the velocity of particle id at iteration k+1, v_{id}^{k} is the velocity of particle id at iteration k. The inertial mass represented by ω has a value between 0.4 and 0.9, and (c1,c2) are the acceleration coefficients (cognitive and social coefficients), r1 and r2 are the random numbers between 0 and 1, x_{id}^{k} is the current position of particle id at the k iteration, $pBest_{id}$ is the best previous position of the id particle, $gBest$ is the position of the best particle in the swarm, and x_{id}^{k+1} is the position of id particle at k+1 iteration. The speed values are between Vmin and Vmax, and N is the size of swarms.

The inertia is calculated by $\omega \times v_{id}^{k}$ and $c_1 r_1 \left(pBest_{id}\text{-}x_{id}^{k}\right)$ represents a memory (the particle is attracted to the best point in its trajectory) and $c_2 r_2 \left(gBest - x_{id}^{k}\right)$ represents the cooperation or information exchange (the particle is attracted to the best point found by all particles).

3.2 Sequential minimal optimization

- Support Vector Machine

Support Vector Machine (SVM) [19] is one of the most popular supervised machine learning algorithms. These is a classification model by evaluating data and identify patterns that retains excellent long generalization capabilities with an integrated resistance to overtraining. This generalization is based on solid theoretical foundations introduced by Vapnik [20]. The basic concept of the SVM regression is to map the non-linearity of data x in a space of high characteristic dimension, and solving a linear regression problem in this feature space.

In the classification of support vector, the separation function is a linear combination of grains as given in equation (3) and are in contact with the support vector,

$$f(x) = \sum_{i \in S} \mu_i \, y_i x_i^t x + b \qquad (3)$$

Where μ_i is a Lagrange xi factor of training models, yi {+ 1, -1} is the corresponding class labels and S denotes the set of support vectors.

- Sequential minimal optimization

The algorithm (SMO) [21] is known as one of the most effective solutions for the training phase of support vector machines. In SMO algorithm, the analytical approach is adopted to solve quadratic programming in order to avoid the complexity of traditional iterative process, and the cumulative error in the process.

SMO quickly solve the problem of quadratic programming (SVM) without additional storage array and without invoking an iterative numeric routine for each sub-problem. SMO down the overall QP problem into several sub-problems similar QP. For a standard SVM, the problem of the smallest possible optimization has two Lagrange multipliers because the Lagrange multipliers must obey a linear equality constraint. At each stage, SMO chooses two Lagrange multipliers to optimize jointly. This is done to find the optimum values for the Lagrange multipliers and SVM is updated to reflect the new optimum values.

First, a Lagrange Multiplier (a1) select violates Karush-Kuhn-Tucker condition for the optimization problem. After that, the second Lagrange multipliers (a2) select and optimize the pair (a2, a2). This process is repeated to obtain the optimum values for these multipliers, and updates the SVM to reflect the new optimal value.

The main advantage of SMO and its analytical approach for solving quadratic programming (QP) that it does not need to store the kernel matrix such as numerical algorithms and greatly reduced space and size of the storage which is quadratic as a function of sample numbers, and it’s very useful in cases of large data sets, as in our case.

4 Architecture of the proposed IDS

In this section, we’ll describe the PSO-SMO system proposed for classification. This study aims first to optimize the accuracy of SMO classifier by detecting the subset of the best information features and estimation of the best values for the regularization kernel parameters of the SMO model. The PSO-SMO algorithm combines two methods of learning, by optimizing the parameters of SMO using PSO.

4.1 Background

We’ll define the proposed Framework and its purpose (Figure 1), then illustrate how to use SMO-PSO algorithm to detect outliers on the network traffic data set.

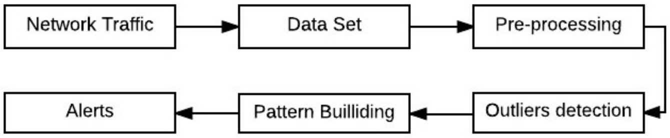

Figure 1: The Framework of the proposed approach

The main theme of data mining approaches is to take a centric data perspective and consider intrusion detection as a process of data analysis [22]. It includes four key steps:

(1) The capture of packets transferred over the network

(2) Pull out an extensive set of features that can describe a network or a host session

(3) Pre-processing data

(4) Acknowledge the model to accurately describe the abnormal and normal activities behavior by applying data mining

(5) Intrusion detection using built model.

In our research, we assume that steps (1) and (2) has been developed and are already available for learning and testing phases. Our work is based around the steps (3), (4) and (5).

IDS sniffs network traffic and built data set in pre-processing. After that, the service-based models are built on the data set using the PSO-SMO. With the built model, we can find outliers relating to each type of attacks. Then, the system will trigger alerts when outliers are detected.

4.2 Data set

Cyber Systems and Technology Group of MIT Lincoln Laboratory [23] simulated LAN US Air Force LAN with multiple attacks and captured nine weeks TCPdump data. This database was first used for competitions kdd99, but since it has become the database test to the IDS's based on a behavioral approach.

Each connection record consists of approximately 100 bytes. This was converted into about 49 * 105 connection vectors each one contains 41 fields.

This database is collected by simulating attacks on different platforms such as Windows, UNIX... Four gigabytes of raw data compressed TCP dump transformed into five million connections files. The attacks are divided into four main categories:

• Denial of Service Attack (DOS): Is a computer attack aimed to cripple a service or to prevent legitimate users of a service to use, the attacker uses different techniques:

-Inundation: A network in order to prevent its operation

-The Disruption of connections between two machines, preventing access to a particular service

-Obstruction access to a service to a particular user

• User to Root Attack (U2R): An exploit in which the attacker starts with access to a regular user account on the system and is able to exploit certain vulnerabilities, he gets administrator access to the system

• Remote to Local Attack (R2L): Occurs when an attacker is able to send packets to a machine on a network without a local account

• Probing Attack is an attempt to gather information on a network with the apparent aim of circumventing its security.

4.3 Optimization of the SMO by PSO

PSO begins with n particles chosen randomly and seeks the iterative optimum particle. Each particle is an M-dimensional vector which represents a candidate's solution. The SMO classifier using the RBF kernel functions (for this faculty analyse high-dimensional data) to built a wide range candidates solution to evaluate its performance until the optimal solution. The PSO algorithms guide the selection of potential subsets that lead to better prediction accuracy. The algorithm uses the most apt particles to contribute to the generation of the candidate's n-particles. Thus, on average, each successive population of candidate particles is better than its predecessor. The technique 10-folds crossover validation is used. The k-folds crossover validation is typically used to reduce the error resulting from random sampling in comparing accuracies of a number of predictive models. The study divided the data into 10-folds where nine for training and one fold for test. This process continues until the performance is satisfactory SMO. PSO is used (Algorithm 1) discovering the best features combinations.

The approach proposed PSO-SMO is described as follows:

Algorithm 1 : Proposed approach

```
Training data set: kdd99_10p
Test data set: kdd99_full
Pre-processing is applied to the Training and Testing data sets
Swarm initialization: positions, velocities of particles, accelerations factors
Split Training data set:  The original Training data is randomly partitioned into K = 10 equal
                          sized subsets
For k=1 to K do
For j = 1 to the number of particles do
Train SMO with particle(j)
Compute accuracy using K-1 subsets;
Validation of SMO model with K(i) subset;
If accuracy(j) > Pbest(j)
          Pbest(j) = accuracy(j);
End if
If  Pbest(j) > Gbest
          Gbest = Pbest(j);
End if
```

```
update the velocity by : $v_{id}^{k+1} = \omega \times v_{id}^{k} + c_1 r_1 (pBest_{id} - x_{id}^{k}) + c_2 r_2 (gBest - x_{id}^{k})$
update particle position using : $x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1}$
End for
End for
Train SMO using all KDD99_10P data set with best particle (subsets)
Evaluate a model using a full kdd99 data set
```

5 EXPERIMENT SETUP AND PERFORMANCE EVALUATION

5.1 Simulation Tools

In this section, we evaluate the performance of the proposed model. All experiments were conducted on a calculation station 24 CPU Intel Core 2.13GHz, 48GB RAM, running under Linux CentOS 7. The implementation was coded using the Java language.

5.2 Pre-processing

In anomaly detection data pre-processing is an important step, it has a key role on the accuracy and the ability of IDS based on anomaly. In our case the data pre-processing is to rely mainly on packet headers knowledge to identify the most relevant parts of the network traffic and the construction of training data.

- Data transformation

The column 2,3 and 4 are types of chains character (String) to the simplified working classifier, which converted all the data in these numeric columns.

- Normalization

In our approach, standardization was performed by scaling numerical characteristics relative to their mean and standard deviation; this prevents features with large numeric values to dominate other features. We also applied a data reduction to the training data.

- Data Reduction

The selection of features is a very active domain in recent years. Its uniqueness is in the context of Data Manning. Indeed, in very large database, it becomes crucial for applications such as IDS, complex industrial processes. This is to summarize and intelligently extract knowledge from raw data to maximally simplify the work of the classification model.

- Features Selection

From an artificial intelligence perspective, create a classifier means creating a template for data, or perfect for a model is to be as simple as possible. Reducing the number of parameters then reduces the number of parameters necessary for the description of this model.

- It improves the classification performance: learning time, his speed and power of generalization.

- It increases the comprehensibility of data. This data selection is to select an optimum subset of relevant variables from a set of original variables.

The KDD99 has 41 variables, it is relatively a large number to be processed by the classifier, the latter in the learning phase cannot complete execution within a reasonable time, then the selection can reduce the feature space. We discover that the values of 7,8,9,11,14,15,17,18,20 and 21 columns are the same. By removing the 10 columns of the classification model the calculation speed also improved and simplified.

5.3 Anomaly Detection Results

This section describes the obtained results from the experiment by applying the proposed algorithms on the data set kdd99. The performance of the proposed method of intrusion detection was evaluated on all KDD99 data set, 10% of the KDD99 data set were used for training the model after the pre-processing step. Table 1 illustrate the confusion matrix, we achieved 97.37% of instances classified correctly and only 2.6292% of false classified instances.

Table 1: Detailed confusion matrix

Actual Class	***Classified Class***					***% Correct***
	Normal	**DOS**	**Probe**	**R2L**	**U2R**	
Normal	858524	56768	57356	45	88	88,25 %
DOS	58	3883112	244	0	0	99,99 %
Probe	423	5088	35532	58	1	86,45 %
R2L	433	2	1	688	2	61,10 %
U2R	20	0	3	1	28	53,85 %

For simplified evaluation of our system, besides the classical accuracy measure, we have used two standard metrics of detection rate and false positive rate developed for network intrusions derived in [24]. Table 2 shows these standard metrics.

Table 2: Standard metrics for system evaluation

		Predicted Label	
		Normal	*Intrusion*
Actual class	*Normal*	TN(858524)	FN(114257)
	Intrusion	FP(934)	TP(3919672)

Figure 2 shows the detection rate classified by attacks, the proposed algorithm has detected 99.99% of DOS attacks whom are the most used by hackers. For Probe attacks a rate of 86.45% is correctly classified, 61.10% for R2L attacks and 53.85% for U2R, the detection rate of R2L is minimal compared with that of the other categories because some of the attacks, exhibit features that are highly similar to those of Normal and may match these features 100% such that they cannot be 100% classified as attacks. The low rate of detection of U2R attacks can be explained by the insufficient number of data learning about these, unlike the DOS attacks and Probe's which are quite numerous and diversified to build a more accurate detection model.

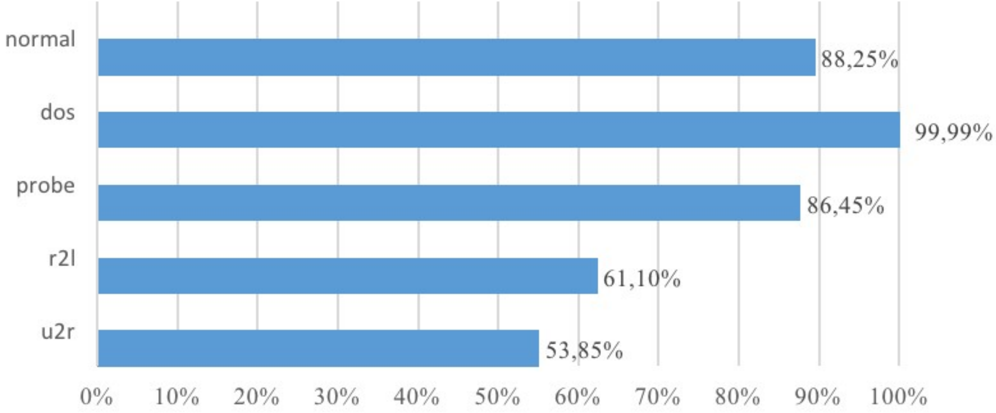

Figure 2: The per attack detection rates

To validate the results of the prediction we compare the proposed technique (SMO + PSO) with SMO classification model (1) only formed with the raw of the kdd99_10p data set and another model SMO (2) too, but trained with submitted data to the pre-treatment step before we have detailed for evaluating the sets of work performance.

Figure 3 compares the results of the different approaches. Note that the results of the model SMO (1) is better than the model SMO (2) because the pre-processing simplified learning classifier and improved the accuracy of the model. The hybrid model (SMO + PSO) is formed by the same training data as SMO (1) provides best detection accuracy by another two percentage data correctly classified (Accuracy) of 97.37%.

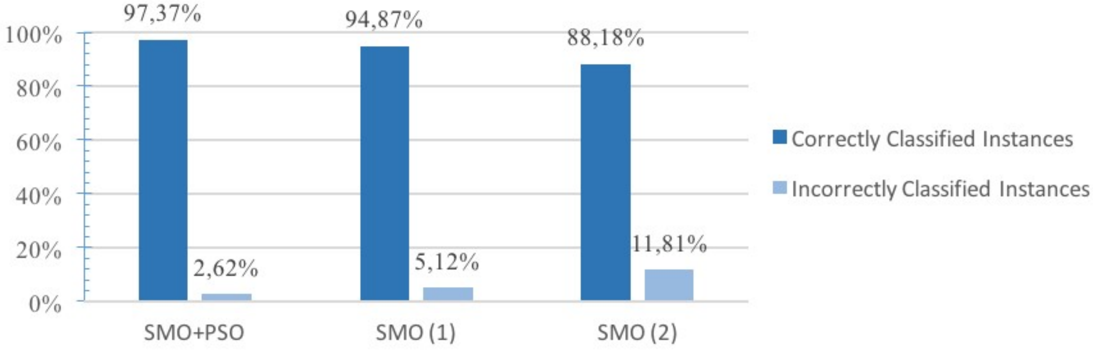

Figure 3: Comparison of results

To compare a detection rate, Figure 4 compares the proposed method with approaches that use only the entire KDD Cup 1999 dataset as a testing dataset because several researchers use only part of the KDD Cup 1999, which may distort the result, since many new attacks in the full dataset have not appeared in the training dataset, and they account for approximately 10% of the KDD99. They cripple the performance of classifiers, and allocating the attacks becomes difficult. The above results show that our approach enhances the performance of IDS. PSO-SMO is more reliable than state-of-the-art methods.

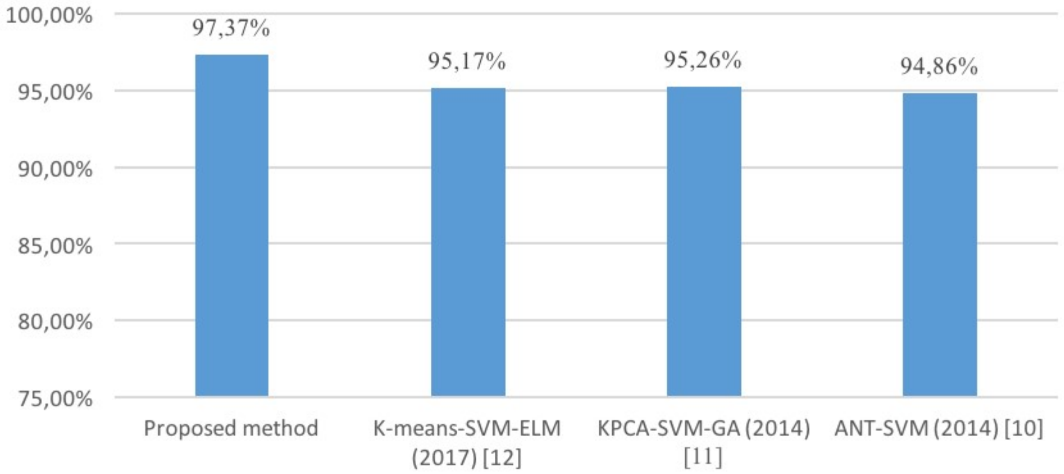

Figure 4: Comparison of proposed model with other methods by detection rate

6. Conclusion and futures Works

The volume of data flowing through today's networks is important that is why it's important to form a robust detection solution, we opted for a behavioral approach. In this paper, a method of applying algorithm SMO coupled with PSO is presented to the network intrusion detection system to effectively detect various types of network intrusions.

To evaluate the performance of the proposed system we have used data from the KDD Cup99, the comparison of experimental results show that the SMO optimized with PSO which offer the best performance with a detection rate of 97.37%.

For future work, we want to develop more approaches to combine several machine learning techniques into one predictive model, using meta-algorithms, to increase the rate of detection of attacks.

Acknowledgements

This work is supported in part by the Superior School of Technology, Moulay Ismail University Meknes, which has provided the calculation station where we have executed our experiences.

References

1. J. P. Anderson. Computer Security Threat Monitoring and Surveillance. Technical Report, James P. Anderson Company, Fort Washington, 1980.
2. D.E. Denning, An Intrusion-Detection Model. IEEE Transactions on Software Engineering,13 (2), 222-232, 1987.
3. L.D. Silva, A.C. Santos, T.D. Mancilha, J. D. Silva, and A. Montes. Detecting attack signatures in the real network traffic with ANNIDA. Expert Systems with Applications. 34 (4), 2326-233, 2008.
5. N. Hubballi and V. Suryanarayanan, False alarm minimization techniques in signature-based intrusion detection systems: A survey, Computer Communications, 49, 1–17, 2014.
6. W. Lee, S.J. Stolfo and K.W. Mok, A data mining framework for building intrusion detection models, proceedings of IEEE Symposium on Security and Privacy. 120–132 (California 1999).
7. M. Thottan and C. Ji, Anomaly detection in IP networks, IEEE Trans. Signal Process. 51 (8), 2191–2204, 1999.
8. S. Song and L. Ling, Flow-based Statistical Aggregation Schemes for Network Anomaly Detection, proceeding of IEEE International Conference on Networking, Sensing and Control. 786 – 791, (Florida ,2006).
9. CSICO Company, Cisco IOS NetFlow http://www.cisco.com/warp/public/732/Tech/nmp/netflow/. (mars, 2017).
10. W. Feng, Q. Zhang, G. Hu, & J. X.Huang, Mining network data for intrusion detection through combining SVMs with ant colony networks, Future Generation Computer Systems, 37, 127–140.
11. F. Kuang, W. Xu, & S. Zhang, A novel hybrid KPCA and SVM with GA model for intrusion detection. Applied Soft Computing Journal, 18, 178–184, 2014.
12. L. A. Wathiq, A. O. Zulaiha, Z. A. Mohd , Multi-Level Hybrid Support Vector Machine and Extreme Learning Machine Based on Modified K-means for Intrusion Detection System, Expert Systems with Applications, 67, 296-303, 2017.
13. A. Saied, R. E. Overill and T. Radzik.Detection of known and unknown DDoS attacks using Artificial, Neural Networks, 172, 385–393, 2016.
14. Monowar H. Bhuyan, Bhattacharyya DK, Kalita JK. An effective unsupervised network anomaly detection method. In proceeding of International conference on advances in computing, communications and informatics, 533-539, (india, 2012).
15. E.M. Knorr, R.T. Ng and V. Tucakov. Distance-based outliers: algorithms and applications, VLDB Journal, 8 (3), 237–253, 2000.
16. G.V. Nadiammai and M. Hemalatha, Effective approach toward Intrusion Detection System using data mining techniques, Egyptian Informatics Journal, 15 (1), 37–50, 2014.
17. J. Kennedy, R.C. Eberhart, Particle swarm optimization, In Proceedings of the IEEE International Conference on Neural Networks, pp. 1942–1948,(Australia 1995).
18. J. Kennedy, R.C. Eberhart. “A discrete binary version of the particle swarm algorithm”, Systems, Man, and Cybernetics, 1997. Computational Cybernetics and Simulation, IEEE International. 4104–4109, (Orlando1997).
19. C. C. Burges, A tutorial on support vector machines for pattern recognition, Data Mining and Knowledge Discovery. 2 (2), 121-167, 1998.
20. C. Cortes, V. Vapnik, "Support vector networks", Machine Learning, 20 (30).273-297, 1995.
21. T. Platt, Sequential minimal optimization: A fast algorithm for training support vector machines, technical report msr-tr-98-I4, Microsoft Research, 1998.
22. W. Lee, S.J. Stolfo, K.W. Mok, Mining audit data to build intrusion detection models, In Proceedings of the 4th International Conference on Knowledge Discovery and Data Mining. 66–72, (new York, 1998).

23. Mahbod Tavallaee, Ebrahim Bagheri, Wei Lu, and Ali A. Ghorbani. “A Detailed Analysis of the KDD CUP 99 Data Set, In Proceedings of the 2009 IEEE Symposium on Computational Intelligence in Security and Defense Applications,1-6, (Ottawa, 2009).
24. M. M. M. Hassan, Current Studies on Intrusion Detection System, Genetic Algorithm and Fuzzy Logic”, International Journal of Distributed and Parallel Systems, 4(2),35-47, 2013.