INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATIONS AND ROBOTICS
ISSN 2320-7345

A SURVEY ON INTRUSION DETECTION SYSTEM BASED SUPPORT VECTOR MACHINE ALGORITHM

Peyman Asgharzadeh[1], Shahram Jamali[2]

[1]*Department of Computer Engineering, Germi Branch, Islamic Azad University, Germi, Iran*
[2] *Department of Computer Engineering, University of Mohaghegh Ardabili , Ardabil, Iran*
Author Correspondence: peyman.asgharzadeh@gmail.com

Abstract: - Whenever an intrusion occurs, the security and value of a computer system is compromised. Network-based attacks make it difficult for legitimate users to access various network services by purposely occupying or sabotaging network resources and services. This can be done by sending large amounts of network traffic, exploiting well-known faults in networking services, and by overloading network hosts. Intrusion Detection attempts to detect computer attacks by examining various data records observed in processes on the network and it is split into two groups, anomaly detection systems and misuse detection systems. Anomaly detection is an attempt to search for malicious behaviour that deviates from established normal patterns. Misuse detection is used to identify intrusions that match known attack scenarios. Our interest here is in anomaly detection and our proposed method is a scalable solution for detecting network-based anomalies. We use Support Vector Machines (SVM) for classification. The SVM is one of the most successful classification algorithms in the data mining area, but its long training time limits its use. Support Vector Machines (SVM) are the classifiers which were originally designed for binary classification. The classification applications can solve multi-class problems. The construction order of binary tree has great influence on the classification performance. In this paper we are studying an algorithm.

Keywords*:* Intrusion detection system, support vector machine.

1. Introduction

Security is becoming a critical issue as the Internet applications are growing. The current security technologies are focusing on encryption, ID, firewall and access control. But all these technologies cannot assure flawless security. The system security can be enhanced by Intrusion detection. The ability of an IDS to classify a large variety of intrusions in real time with accurate results is important. The patterns of user activities and audit records are examined and the intrusions are located. IDSs are classified, based on their functionality, as misuse detectors and anomaly detectors. Misuse detection system uses well defined patterns of attack which are matched against user behaviour to detect intrusions. Usually, misuse detection is simpler than anomaly detection as it uses rule based or signature comparison methods. Anomaly detection requires storage of normal usage behaviour and operates upon audit data generated by the operating system. Support vector machines (SVM) are the classifiers which were originally designed for binary classification, can be used to classify the attacks. If binary SVMs are combined with decision trees, we can have multiclass SVMs, which can classify the four types of attacks, Probing, DoS, U2R, R2L attacks and Normal data, and can prepare five classes for anomaly detection. Our aim is to improve the training time, testing time and accuracy of IDS using the hybrid approach.
Since the tragic events of September 11, 2001, insuring the integrity of computer networks, both in relation to security and with regard to the institutional life of the nation in general is a growing concern. Security and

defence networks, proprietary research, intellectual property, data based market mechanisms which depend on unimpeded and undistorted access can all be severely compromised by intrusions.

We need to find the best way to protect these systems. An intrusion can be defined as “any set of actions

That attempts to compromise the integrity, confidentiality, or availability of a resource”. User authentication (e.g., using passwords or biometrics), avoiding programming errors, and information protection (e.g., encryption) have all been used to protect computer systems. As systems become more complex, there are always exploitable weaknesses due to design and programming errors, or through the use of various “socially engineered” penetration techniques. For example, exploitable “buffer overflow” still exists in some recent system software because of programming errors. Elements central to intrusion detection are resources to be protected in a target system, i.e., user accounts, file systems, system kernels, etc.; models that characterize the “normal” or “legitimate” behaviour of these resources; techniques that compare the actual system activities with the established models identifying those that are “abnormal” or “intrusive”. In pursuit of a secure system, different measures of system behaviour have been proposed, on the basis of an ad hoc presumption that normalcy and anomaly (or illegitimacy) will be accurately manifested in the chosen set of system features. Intrusion Detection attempts to detect computer attacks by examining various data records observed through processes on the same network. These attacks are split into two categories, host-based attacks and network-based attacks. Host-based attacks target a machine and try to gain access to privileged services or resources on that machine. Host-based detection usually uses routines that obtain system call data from an audit-process which tracks all system calls made on behalf of each user. Network-based attacks make it difficult for legitimate users to access various network services by purposely occupying or sabotaging network resources and services. This can be done by sending large amounts of network traffic, exploiting well-known faults in networking services, overloading network hosts, etc. Network-based attack detection uses network traffic data (i.e., tcpdump) to look at traffic addressed to the machines being monitored. Intrusion detection systems are split into two groups, anomaly detection systems and misuse detection systems. Anomaly detection is the attempt to identify malicious traffic based on deviations from established normal network traffic patterns. Misuse detection is the ability to identify intrusions based on a known pattern for the malicious activity. These known patterns are referred to as signatures. Anomaly detection is capable of catching new attacks. However, new legitimate behaviour can also be falsely identified as an attack, resulting in a false positive. Our research will focus on network level systems. The problem with current state-of-the-art is to reduce false negative and false positive rate (i.e., we wish to minimize “abnormal normal” behaviour). At the same time, a real-time intrusion detection system should be considered. It is difficult to achieve both. The SVM is one of the most successful classification algorithms in the data mining area, but its long training time limits its use. Many applications, such as Data Mining and Bio-Informatics, require the processing of huge data sets. The training time of SVM is a serious obstacle in the processing of such data sets. According to, it would take years to train SVM on a data set consisting of one million records. Many proposals have been submitted to enhance SVM in order to increase its training performance, either through random selection or approximation of the marginal classifier. However, such approaches are still not feasible with large data sets where even multiple scans of entire data set are too expensive to perform, or result in the loss through over-simplification of any benefit to be gained through the use of SVM. This paper proposes a new approach for enhancing the training process of SVM when dealing with large training data sets. It is based on the combination of SVM and clustering analysis. The idea is as follows: SVM computes the maximal margin separating data points; hence, only those patterns closest to the margin can affect the computations of that margin, while other points can be discarded without affecting the final result. Those points lying close to the margin are called support vectors (see Sect. 3 for more details). We try to approximate these points by applying clustering analysis. In general, using hierarchical clustering analysis based on dynamically growing self-organizing tree (DGSOT).involves expensive computations, especially if the set of training data is large. However, in our approach, we control the growth of the hierarchical tree by allowing tree nodes (support vector nodes) close to the marginal area to grow, while halting distant ones. Therefore, the computations of SVM and further clustering analysis will be reduced dramatically. Also, to avoid the cost of computations involved in clustering analysis, we train SVM on the nodes of the tree after each phase/iteration, in which few nodes are added to the tree. Each iteration involves growing the hierarchical tree by adding new children to the tree. This could cause a degradation of the accuracy of the resulting classifier. However, we use the support vector set as a priori knowledge to instruct the clustering algorithm to grow support vector nodes and to stop growing non-support vector nodes. By applying this procedure, the accuracy of the classifier improves and the size of the training set is kept to a minimum. We report results here with one benchmark data set, the 1998 DARPA. Also, we compare our approaches with the Rocchio Bundling algorithm, recently proposed for classifying documents by reducing the number of data points. Note that the Rocchio Bundling method reduces the number of data points before feeding those data points as support vectors to SVM for training. On the other hand, our clustering approaches intertwined with SVM. We have observed that our approaches outperform Pure SVM and the Rocchio Bundling technique in terms of accuracy, false positive (FP) rate, false negative (FN) rate, and processing time. The main contributions of this work are as follows: First, to reduce the training time of SVM, we propose a new support vector selection technique using clustering analysis. Here, we combine the clustering analysis and SVM training phases. Second, we show analytically the degree to which our approach is asymptomatically quicker than pure SVM, and validate this claim with experimental results. Finally, we compare our approaches with random selection, and Rocchio Bundling on a benchmark data set, and demonstrate impressive results in terms of training time, FP rate, FN rate, and accuracy.

2. Related Work

Here, first, we present related work relevant to intrusion detection, and next, we present related work for the reduction of training time of SVM. In particular, we will present various clustering techniques as data reduction mechanisms. With regard to intrusion detection, as noted earlier, there are two different approaches to intrusion detection system (IDS): misuse detection and anomaly detection. Misuse detection is the ability to identify intrusions based on a known pattern for the malicious activity. These known patterns are referred to as signatures. The second approach, anomaly detection, is the attempt to identify malicious traffic based on deviations from established normal network traffic patterns. “A State Transition Analysis Tool for Intrusion Detection” (STAT) and “Intrusion Detection in Our Time” (IDIOT) are misuse detection systems that use the signatures of known attacks. Lee et al. propose a data mining framework for intrusion detection which is misuse detection. Their goal is to automatically generate misuse detection signatures from classified network traffic. Anomaly detection is capable of catching new attacks. However, new legitimate behaviour can also be falsely identified as an attack, resulting in a false positive. In recent years, there have been several learning-based or data mining-based research efforts in intrusion detection. Instead of network level data, researchers may also concentrate on user command-level data. For example, “Anomaly-Based Data Mining for Intrusions,” ADMIT is a user profile dependent, temporal sequence clustering based real-time intrusion detection system with host-based data collection and processing. In this effort “Next Generation Intrusion Detection Expert System”, NIDES and “Event Monitoring Enabling Responses to Anomalous Live Disturbances,” EMERALD create user profile based on statistical method. A few other groups advocate the use of neural networks in intrusion detection. Some of them rely on a keyword count for a misuse detection system, along with neural networks. Attack specific keyword counts in network traffic are fed as neural network input. Ghosh et al. use a neural network to extract program behaviour profiles instead of user behaviour profiles, and later compare these with the current system behaviour. Self-organizing maps (SOM) and support vector machine have also been used as anomaly intrusion detectors. An SOM is used to cluster and then graphically display the network data for the user to determine which clusters contained attacks. SVM is also used for an intrusion detection system. Wang et al. use “one class SVM” based on one set of examples belonging to a particular class and no negative examples rather than using positive and negative examples. Neither of these approaches addresses the reduction of the training time of SVM, which is what prohibits real-time usage of these approaches. With regard to the training time of SVM, random sampling has been used to enhance the training of SVM. Sub-sampling speeds up a classifier by randomly removing training points. Balacazar et al. use random sampling successfully to train SVM in many applications in the data mining field. Shih et al. use sub-sampling in classification using a Rocchio Algorithm along with other data reduction techniques. Sub-sampling surprisingly has led to an accurate classification in their experiments with several data sets. However, Yu et al. show that random sampling could hurt the training process of SVM, especially when the probability distribution of training and testing data were different. Yu et al. use the idea of clustering, using BIRCH, to fit a huge data set in the computer memory and train SVM on the tree’s nodes. The training process of SVM starts at the end of building the hierarchical tree causing expensive computations, especially when the data cannot fit in the computer memory or the data is not distributed uniformly. The main objective of the clustering algorithm is to reduce the expensive disk access cost; on the other hand, our main focus is to approximate support vectors in advance. Furthermore, our use of clustering analysis goes in parallel with training SVM, i.e., we do not wait until the end of building the hierarchical tree in order to start training SVM. It is a legitimate question to ask why we use hierarchical clustering rather than partitioning/flat clustering (e.g., K-means). Partitioning/flat clustering directly seeks a partition of the data which optimizes a predefined numerical measure. In partitioning clustering, the number of clusters is predefined, and determining the optimal number of clusters may involve more computational cost than that of the clustering itself. Furthermore, a priori knowledge may be necessary for initialization and the avoidance of local minima. Hierarchical clustering, on the other hand, does not require a predefined number of clusters or a priori knowledge. Hence, we favour hierarchical clustering over flat clustering. Hierarchical clustering employs a process of successively merging smaller clusters into larger ones (agglomerative, bottom-up), or successively splitting larger clusters (divisive, top down). Agglomerative algorithms are more expensive than divisive algorithms and, since we need a clustering algorithm that grows in a top-down fashion and is computationally less expensive, we favour the use of divisive hierarchal clustering over agglomerative algorithms.

3. Support Vector Machine

Binary classification problems can be solved using SVM. An SVM maps linear algorithms into non-linear space. It uses a feature called, kernel function, for this mapping. Kernel functions like polynomial, radial basis function are used to divide the feature space by constructing a hyper-plane. The kernel functions can be used at the time of training of the classifiers which selects support vectors along the surface of this function. SVM classify data by using these support vectors that outline the hyper-plane in the feature space. This process will involve a quadratic programming problem, and this will get a global optimal solution. Suppose we have N training data points {(x1, y1), (x2,y2), (x3, y3), ..., (xN , yN)}, where xi € Rd and yi €{+1,−1}. Consider a hyper-plane defined by (w, b), where w is a weight vector and b is a bias. The classification of a new object x is done with

$$f(x) = \text{sign}(w.x + b) = \text{sign}(\sum_{i}^{N} \alpha_i \, y_i (x_i \cdot x) + b) \qquad (3\text{-}1)$$

The training vectors xi occurs only in the form of a dot product. For each training point, there is a Lagrangian multiplier αi. The Lagrangian multiplier values αi reflects the importance of each data point. When the maximal margin hyper-plane is found, only points that lie closest to the hyper-plane will have αi > 0 and these points are called support vectors. All other points will have αi = 0. That means only those points that lie closest to the hyper-plane, give the representation of the hypothesis/classifier. These data points serve as support vectors. Their values can be used to give an independent boundary with regard to the reliability of the hypothesis/classifier.

3.1 Multiclass Support Vector Machine

Multiclass SVM constructs k different classes at the training phase of IDS. Some typical methods [2][10] , for construction of multiclass SVM are one-versus-rest(OVR), one-versus one(OVO) and method based on Directed Acyclic Graph(DAG).

3.1.1 One-versus- Rest

It constructs k two-class SVMs. The ith SVM (i = 1, 2, . . . , k) is trained with all training patterns. The ith class patterns are labelled by 1 and the rest patterns are labeled by −1. The class of an unknown pattern x is determined by argument maxi=1,2,...,k fi(x), where fi(x) is the decision function of the ith two-class SVM. In short, a binary classifier is constructed to separate instances of class yi from the rest of the classes. The training and test phase of this method are usually very slow.

3.1.2 One-versus-one

It constructs all possible two-class SVM classifiers. There are k(k − 1)/2 classifiers by training each classifier on only two out of k classes. A Max Wins algorithm is used in test phase,: each classifier casts one vote for its favored class, and finally the class with most votes wins. The number of the classifiers is increased super linearly when k grows. OVO becomes slow on the problem where the number of classes is large.

4. Intrusion Detection System

Intrusion Detection System (IDS) has quickly established as the most important element of security infrastructure. Intrusion is an attempted act of using computer system resources without privileges, causing incidental damage. Intrusion Detection is the process of monitoring the events occurring in a computer system or network and analyzing them to sign in possible incidents. An ID monitors network traffic, monitoring the events occurring in a computer system or network and analyzing them for sign in possible incidents. If it detects any threat then alerts the system or network administrator. There are two performance evaluation variables criteria, Detection Rate (DR) which is defined as the ratio of number of correctly detected attacks to the total number of attacks & False alarm rate(FAR) which is ratio of the number of normal connection that are misclassified as attacks to total number of normal connections. Intrusion Detection must be able to identify intrusion with high accuracy and it must not confuse normal action with the occurrence of a system with intrusive ones. Construction of efficient Intrusion Detection is a challenging task so it must have a high attack Detection Rate (DR) with low False alarm rate (FAR) at the same time.

4.1 Type of Intrusion Detection System

Intrusion Detection System is broadly classified on the basis of following two Criteria:

Based on Data Collection mechanism
Based on Detection Techniques

On the basis of data collection mechanisms IDS is categorized into Host-Based Intrusion Detection System (HIDS) and Network-Based Intrusion Detection System (NIDS). Host-based IDS is dependent for support on capturing local network traffic to the specific host. This local host analyzes and process data which is used to secure the activities of this host and informs about the attacks in the network. HIDS analysis events mainly related to OS information. Network-based intrusion detection system (NIDS) works on network and observes the network traffic. NIDS analyses network related traffic volumes, IP address service port etc. which are able to detect attack from outside, examine packet header and entire packet. On the basis of detection techniques IDS is classified as Misuse Detection and Anomaly Detection. Misuse Detection it involves searching network traffic for a series of malicious activity within the analyzed data. The main advantage of this technique is that it provides very good detection results for specified, well known attacks & is very easy to develop and understand. However they are not capable of detecting novel attacks. Anomaly intrusion detection system (AIDS) uses normal usage behavior patterns to identify the intrusion. The normal usage patterns are constructed from the statistical measures of the system features. While the anomaly behavior detecting system generates a standard traffic sketch & employs it to detect any abnormal traffic pattern and attempts of intrusion. The three main vital factor's that impact the quality anomaly detection is Feature Selection, Data value normalization and Classification technique. According to the type of processing related to the ''Behavioral'' model of the target system, Anomaly Detection Techniques can be classified into three main categories: Statistical based, Knowledge-based, and Machine Learning based. In the Statistical based, the behavior of the system is represented by a random view point. On the other hand, Knowledge-based Anomaly network intrusion detection techniques try to capture the claimed behavior from available system data (protocol specifications, network traffic instances, etc.). Finally Machine learning techniques are based on establishing an explicit or implicit model that enables the patterns analyzed to be categorized. The comparison of all the three AIDS as shown in Table 1. As our research is on Machine learning based Anomaly detection system so in the next section a short introduction of Machine Learning Techniques used in Anomaly Intrusion Detection System is described.

Table 1. Comparison of All the Three AIDS

Technique	**Advantages**	**Disadvantage**
Statistical-based:- stochastic behavior	*Future knowledge about Normal activity is not required. Exact and accurate notification About intruder's activities.*	*Parameters and metrics are very difficult to set. Easily influenced could be trained by attackers.*
Knowledge-based:-	*Robustness. Flexibility and Scalability.*	*Difficult and time-consuming availability for high-quality Knowledge/data.*
Machine learning based:-	*Flexibility and adaptability. Capture of interdependencies*	*High dependency on the assumption about the behavior accepted for the System. High resource Consuming.*

4.2 Learning Techniques

Learning or training is a process by means of which a network adapts itself to a stimulus by making proper parameter adjustments resulting in production of desired responses. The learning technique can generally classified into two categories as unsupervised learning and Supervised learning. Unsupervised algorithm seeks out similarities between pieces of data in order to determine whether they can be characterized as forming a group. These groups are termed clusters, and there are whole families of clustering machine learning techniques.

In unsupervised classification, often known as 'cluster analysis' the machine is not told how the texts are grouped. Example of unsupervised learning is the self-organizing map (SOM) and Adaptive Resonance Theory (ART). Supervised Machine Learning is the search for algorithms that reason from externally supplied instances to produce general hypotheses, which then make predictions about future instances. In other words, the goal of supervised learning is to build a concise model of the distribution of class labels in terms of predictor features. The resulting classifier is then used to assign class labels to the testing instances where the values of the predictor features are known, but the value of the class label is unknown. Examples of supervised learning are Rough Set, Support Vector Machine and Neural Network.

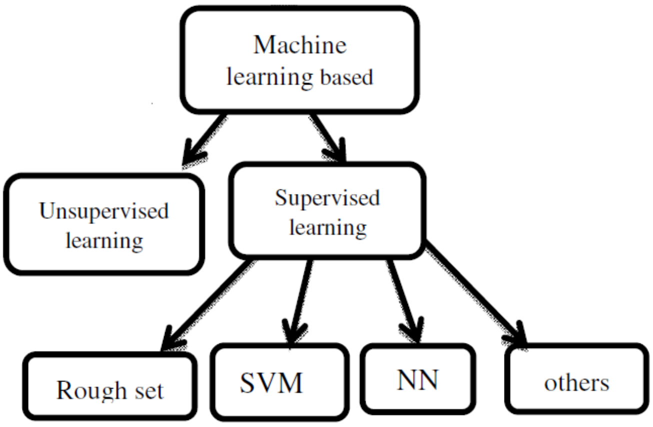

Fig1. Instruction Detection Technique

5. Using Support Vector Machine

The Support Vector Machine (SVM) is a supervised learning method from the field of machine learning applied to both classification and regression based on statistical learning theory. It can find a solution by making a nonlinear transformation of the original input space into a high dimensional feature space where an optimal separating hyper plane can be found, which means that a maximal margin classifier in relation to the training data set can be obtained. Support Vector Machine is effective in reducing the number of alerts, false positive, false negative better, parameter optimization in SVM is very important for its efficiency. A number of methods, such as grid search &evolutionary algorithms have been utilized to optimize the model parameters of SVM. So this section discusses the use of PSO for feature selection & parameter optimization in Network anomaly detection system for SVM. In Tu. Chung et.al Particle Swarm Optimization (PSO) is used to implement a feature selection, and then fitness values are evaluated with a Support Vector Machines (SVMs) which was combined with one-versus rest method for five classification problem. The Binary Particle swarm Optimization (BPSO) is used to serve as feature selection for classification problem. It helps to improve the performance owing to its smaller number of simple parameter settings. Kernel Adatron (KA) SVM is used to evaluate the fitness values of the PSO, which can be obtained by comparing the characteristic of the general test data .Experimental results show that proposed method simplified feature selection and the total number of parameters needed effectively, thereby obtaining a higher classification accuracy compared to other feature selection methods. In Ma.Jing et al. a New method of hybrid Intrusion Detection based on hybridization of Binary Particle Swarm Optimization (BPSO) and Support Vector Machine (SVM).Method is proposed for simultaneous feature selection and parameter optimization. In this combinatorial technique, parameters of SVM and dataset features are represented by every particle position (i.e., a binary series). The modified BPSO is used to obtain the best particle position quickly throughout the search space, which cooperates with SVM for evaluating the fitness of the corresponding particle. Consequently, the optimum features and parameters are chosen at the same time. The main purpose is to find out better parameters for SVM and a feature subset involving key features of network intrusion attacks based on the improved BPSO-SVM. Experimental results show that technique will be useful to reduce the data quantity of large scale dataset and improve the classification ability of the classifier in IDS. In Zhang et al. presents a Hybrid Quantum Binary Particle Swarm Optimization (QBPSO)-SVM based network intrusion wrapper algorithm.. In QPSO each bit of particle is represented by quabit, which has two basic state‘0’ and ‘1’.The quantum superposition characteristic can make a single particle represent several states, thus potentially increases population diversity. The probability representation makes particle mutate according a certain probability to avoid local optimal. When experimented

with the classical intrusion feature selection, it was found that there exist correlation relationship among network intrusion features, so Modified QBPSO based wrapper feature selection is superior to those classical intrusion feature selection methods. The paper reported that, the proposed method is an effective and efficient way for feature selection and detection when tested on the data sets of KDD cup 99. New design of IDS was proposed in Zhou .J et al. which presents optimal selection approach of the SVM parameters based on Particle Swarm Optimization algorithm. PSO parameters selection method not only to ensure that SVM learning ability but also to some extent, improved the generalization ability of SVM and performance of support vector machine classifier. The experimental result shows Particle Swarm Optimization and Support Vector Machine are effective in reducing the number of alerts, false positive, false negative better. In Wang J et al. Simple Particle Swarm Optimization (SPSO) is used to optimize the SVM model parameters and feature selection for IDS. Support vector machine (SVM) has been employed to provide potential solutions for the IDS problem. Firstly feature selection algorithm select important features, and then built intrusion detection systems using these selected features. The training data set is then separated into attack data sets and normal datasets, which are then subsequently, fed into the hybrid PSO-SVM algorithms. Experiment results show that proposed method is not only able to achieve the process of selecting important features but also to yield high detection rates for IDS..
Since efficient classification algorithms are extremely important for intrusion detection, a large number of studies have been conducted. K-Nearest neighbour (KNN) is an extremely simple yet surprisingly effective method for a classification. Its advantages stem from the fact that its decision surface is nonlinear with only a single integer parameter. More importantly, these advantages do not cause the over-fitting (Denning), and it is not restricted to any specific data distribution. The main features of Distance-based classification are summarized as follows.

6. Conclusion

Intrusion Detection based upon Particle Swarm Optimization is currently attracting considerable interest from the research community, being able to satisfy the growing demand of reliable and intelligent Intrusion Detection Systems. The main advantage of PSO is that it is easy to implement & only a few input parameters are needed to be adjusted & is effective in nonlinear optimization problem. Also updation of velocity and position in Particle Swarm Optimization is based on simple equations so it can be efficiently used on large data sets. From the survey done in this paper it is revealed that there are several factors that affects the performance IDS. First is selection & extraction of relevant features. If all features are evaluated then it degrade the IDS performance, so to enhance the performance researchers uses several Supervised Machine Learning techniques each of which has its own pros and cons. Also it has been proven that there is no single generic classifier available that can classify all the attack types effectively so hybridization of different Supervised Machine Learning techniques is done by several researchers. Since the single article cannot be a complete review of the research done in the mentioned area, so only hybridization of PSO with Rough-Set, ANN and SVM & some of the other Machine Learning techniques is discussed here. In this paper, the contributions of research work done in recent years, in each method were summarized and existing research challenges are also defined. It is hoped that this survey can serve as a useful guide for the researchers interested in Particle Swarm Optimization Based Machine learning Oriented Anomaly Network Intrusion Detection System.

REFERENCES

[1] Denning, D.: An intrusion detection model. IEEE Transactions of Software Engineering 13(2), 222–232 (1987)

[2] Lazarevic, A., Kumar, V., Srivastava, J.: Intrusion detection: a survey. In: Managing Cyber Threats: Issues, Approaches, and Challenges, p. 330. Springer (2005)

[3] Garcia-Teodoroa, P., Diaz-Verdejoa, J., Macia-Fernandez, G., Vazquez, E.: Anomaly based network intrusion detection; technique, systems and challenges. Computers and Security 28, 18–28 (2009)

[4] Kennedy, J., Eberhart, R.C.: Particle Swarm Optimization. In: Proceedings of the IEEE International Joint Conference on Neural Networks, pp. 1942–1948 (1995)

[5] Zainal, A., Maarof, M.A., Shamsuddin, S.M.: Feature Selection Using Rough Set in Intrusion Detection. In: IEEE TENCON 2006, Hongkong, November 14-17 (2006)

[6] Zainal, A., Maarof, M.A., Shamsuddin, S.M.: Feature Selection Using Rough-DPSO in Anomaly Intrusion Detection. In: Gervasi, O., Gavrilova, M.L. (eds.) ICCSA 2007, Part I. LNCS, vol. 4705, pp. 512–524. Springer, Heidelberg (2007)

[7] Tian, W., Liu, J.: Network Intrusion Detection Analysis with Neural Network and Particle Swarm Optimization Algorithm. In: 2010 Chinese IEEE Control and Decision Conference, CCDC, pp. 1749–1752 (2010)

[8] Liu, H., Jian, Y., Liu, S.: A New Intelligent Intrusion Detection Method Based on Attribute Reduction and Parameters Optimization of SVM. In: Proceedings of the Second International Workshop on Education Technology and Computer Science (ETCS), pp. 202–205 (2010)

[9] Wang, H.-B., Fu, D.-S.: An Intrusion Detection System Model Based on Particle Swarm Reduction. In: Proceedings of 4th the IEEE International Conference on Genetic and Evolutionary Computing, pp. 383–385 (2010)

[10] Liu, L.-L., Liu, Y.: MQPSO based on wavelet neural network for network anomaly detection. In: Proceedings of the 5th International Conference on Wireless Communications, Networking and Mobile Computing (WiCom 2009), pp. 1 5 (2009)

[11] Liu, Y., Ruhui, M.A.: Wavelet Neural Networks Optimized by QPSO for Network Anomaly Detection. Journal of Computational Information Systems 7(7), 2452–2460 (2011)

[12] Liu, Y.: Wavelet fuzzy neural network based on modified QPSO for network anomaly detection. Applied Mechanics and Materials 20-23, 1378–1384 (2010) 452 K. Satpute et al.

[13] Chen, Z., Qian, P., Chen, Z.: Application of PSO-RBF neural network in network intrusion detection. In: Proceedings of the 3rd International Symposium on Intelligent Information Technology Application, pp. 362–364 (2009)

[14] Liu, Y.: QPSO-optimized RBF Neural Network for Network Anomaly Detection. Journal of Information & Computational Science 8(9), 1479–1485 (2011)

[15] Xu, R., Rui, A., Xiao, F.: Research Intrusion Detection Based PSO-RBF Classifier. In: Proceeding of IEEE 2nd International Conference on Software Engineering and Service Science (ICSESS), pp. 104–107 (2011)

[16] Tu, C.-J., Li-Yeh, C., Jun, Y., Cheng, H.: Feature Selection using PSO-SVM. IAENG International Journal of Computer Science 33(1), IJCS_33_1_18 (2007)

[17] Ma, J., Liu, X., Liu, S.: A New Intrusion Detection Method Based on BPSO-SVM. In: Proceedings of the International Symposium on Computational Intelligence and Design, pp. 473–477 (2008a)

[18] Zhang, H., Gao, H.-H., Wang, X.Y.: Quantum Particle swarm optimization based network Intrusion feature selection and Detection. In: Proceedings of the 17th World Congress The International Federation of Automatic Control, Seoul, Korea (2008)

[19] Zhou, T., Li, Y., Li, J.: Research on intrusion detection of SVM based on PSO. In: Proceedings of the International Conference on Machine Learning and Cybernetics, pp. 1205–1209 (2009)

[20] Wang, J., Hong, X., Ren, R.-R., Li, T.-H.: A Real-time Intrusion Detection System based on PSO-SVM. In: Proceedings of the International Workshop on Information Security and Application (IWISA 2009), pp. 319–321 (2009)

[21] Jun GUO, Norikazu Takahashi, Wenxin Hu. An Efficient Algorithm for Multi-class Support Vector Machines. IEEE- 2008.

[22] Latifur Khan, Mamoun Awad, Bhavani Thuraisingham. A new intrusion detection system using support vector machines and hierarchical clustering. The VLDB Journal DOI 10.1007/s00778-006-0002 , 2007.

[23] V. N. Vapnik. The nature of statistical learning theory. Springer-Verlag,New York. NY, 1995.

[24] Xiaodan Wang, Zhaohui Shi, Chongming Wu and Wei Wang. An Improved Algorithm for Decision-Tree-Based SVM. IEEE-2006.

[25] Pang-Ning Tan, Michael Steinbach, Vipin Kumar. Introduction to data mining. Pearson Education.

[26] K. Crammer and Y. Singer. On the algorithmic implementation of multiclass kernel-based vector machines. Journal of Machine Learning Research, 2:265–292, 2001.

[27] YMahbod Tavallaee, Ebrahim Bagheri, Wei Lu, and Ali A. Ghorbani. A detailed analysis of KDD CUP’99 data set. IEEE-2009.

[28] http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

[29] C. W. Hsu, C. J. Lin. A comparison of methods for multiclass support vector machines. IEEE Trans. On Neural Networks, vol. 13, no. 2, pp.415-425, 2002.

[30] Snehal Mulay, P.R. Devale, G.V. Garje. Decision Tree based Support Vector Machine for Intrusion Detection. ICNIT- 2010, unpublished.

[31] Lili Cheng, Jianpei Zhang, Jing Yang, Jun Ma. An improved Hierarchical Multi-Class Support Vector Machine with Binary Tree Architecture" 978-0-7695-3112- 0/08 2008 IEEE DOI 10.1109/ICICSE.2008

[32] K. Hirasawa, M. Okubo, H. Katagiri, J. Hu, and J. Murata. "Comparison Between Genetic Network Programming (GNP) and Genetic Programming (GP)." *Congress on Evolutionary Computation, pages 1276{1282, 2001*. 2001. 1276-1282.

[33] K.Hwang, Y.Chen, and H.Liu. "Defending Distributed Computing Systems from Malicious Intrusions and Network Anomalies." *In Proc. of the 19th IEEE International Parallel and Distributed Processing Symposium*. 2005.

[34] K.Hwang, Y.Kwok, S.Song, M.Cai, Y.Chen, and Y.Chen. "DHT-Based Security Infrastructure for Trusted Internet and Grid Computing." *International Journal Of Critical Infrastructures* 4 (2006): 412-433.

[35] L. Portnoy, E. Eskin, and S. Stolfo. "Intrusion Detection with Unlabeled Data using Clustering." *ACM Workshop Data Mining Applied to Security(DMSA)*. 2001.

[36] M.Bishop. "Computer Security in the Future." *The ISC International Journal of Information* 3 (2011): 3-27.

[37] M.Kryszkiewicz. "Representative Association Rules and Minimum Condition Maximum Consequence Association Rules." *Lecture Notes in Computer Science*. 1998. 361-369.

[38] M.Park, A.Patcha and J. "An Overview of Anomaly Detection Techniques: Existing Scheme for Host-based Anomaly Intrusion Detection." *Network, IEEE* 23 (2009): 42-47.

[39] N. B. Amor, S. Benferhat, and Z. Elouedi. "Naive Bayes vs Decision Tree in Intrusion Detection." *2004 ACM symposium on Applied computing*. 2004.

[40] N. Lu, S. Mabu, and K. Hirasawa. "Integrated Rule Mining based on Fuzzy GNP and Probabilistic Classification for Intrusion Detection." *Journal of Advanced Computational Intelligence and Intelligent Informatics* 15 (2011).

[41] N. Lu, S. Mabu, T. Wang, and K. Hirasawa. "Distance-based Classification using Average Matching Degree and its Application to Intrusion Detection Systems." *IEEJ Transactions on Electronics, Information and Systems* 132 (2012).

[42] N. Ye, S. M. Emran, X. Li, and Q. Chen. "Statistical Process Control for Computer Intrusion Detection." *DARPA Information Survivability Conference and Exposition*. 2001. 3-14.

[43] N.Jaisankar, SGP.Yogesh, A.Kannan and K.Anand, Intelligent. "Agent Based Intrusion Detection System Using Fuzzy Rough Set Based Outlier Detection." *Soft Computing Techniques* (2012): 147-153.

[44] O.Depren, M.Topallar, E.Anarim, and M.K.Ciliz. "An Intelligent Intrusion Detection System (IDS) for Anomaly and Misuse detection in Computer Networks." *Expert Systems with Applications* 29 (2005): 713-722.

[45] Okamoto, A. Kanaoka and E. "Multivariate Statistical Analysis of Network Traffic for Intrusion Detection." *14th International Workshop on Database and Expert Systems Applications*. 2003. 472-476.

[46] P. D. Williams, K. P. Anchor, J. L. Bebo, G. H. Gunsch, and G. D. Lamont. "CDIS: Towards a Computer Immune System for Detecting Network Intrusions." *4th International Symposium on Recent Advances in Intrusion Detection*. 2001. 117-133.

[47] P. Laskov, P. Dssel, C. Schfer, and K. Rieck. "Learning Intrusion Detection: Supervised or Unsupervised?" *Image Analysis and Processing ICIAP 2005 Lecture Notes in Computer Science*. 2005. 50-57.

[48] H. Kanani, Sh. Jamali, "A Survay on Intelligence Intrusion Detection System" International Journal of Research in Computer Application and Robotics, 2015, 2320-7345.