IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 10, December 2014.
www.ijiset.com
ISSN 2348 – 7968

A novel intrusion detection system based on KPCA and RVM with PSO model

S.Suganya, R.Kavitha, M.E

Infant Jesus College Of Engineeting And Technology

Abstract:

The aim of the present work was to design and develop of a Data mining based Network Intrusion Detection System which can detect intrusions based on misuse detection technique and learning algorithm. The work also aimed at reducing number of false alarms by characterizing the target network with appropriate network parameters and analyzing them with mathematical models. This project proposed the KPCA and RVM with PSO for intrusion detection system. The Relevance Vector Machine (RVM) is a machine learning technique that uses Bayesian inference to obtain parsimonious solutions for regression and probabilistic classification. The RVM has an identical functional form to the support vector machine, but provides probabilistic classification.

1. Introduction:

The idea of detecting the intrusions or system misuses by looking at some kind malicious patterns in the network or user activity was initially conceived by James Anderson in his report titled “Computer Security Threat Monitoring and Surveillance” to US Air Force in the year 1980. In the year 1984, the first prototype of Intrusion Detection System which monitors the user activities, named “Intrusion Detection Expert System” (IDES) was developed. In the year 1988, “Haystack” became the first IDS to use patterns and statistical analysis for detecting malicious activities, but it lacked the capabilities of real time analysis. Meanwhile, there were other significant advances occurring at University of California Davis' Lawrence Livermore Laboratories. In the year 1989, they built an IDS called “Network System Monitor” (NSM) for analyzing the network traffic. This project was subsequently developed into IDS named “Distributed Intrusion Detection System” (DIDS). “Stalker” based on DIDS became the first commercially available IDS and influenced the growth and trends of future IDS. In the Mid 90’s, SAIC developed “Computer Misuse Detection System” (CMDS), a host based IDS. US Air Force’s Cryptographic support Centre developed “Automated Security Incident Measurement” (ASIM), which addressed the issues like scalability and portability. The Network IDS has to operate transparently to avoid the intruders from targeting the IDS itself. So generally the IDS is configured to work in a special mode called “Stealth mode”. In this arrangement, the IDS sniffing interface is put in promiscuous mode without assigning the IP address, thus only listening to the packets flowing across the network keeping its presence transparent from network users. Usually the IDS has two Network interfaces, one to monitor the network and the second one for administrative purposes, like configuring IDS, updating signatures, communication with IDS sensors/Manager ,dispatching alerts etc. Attacker can easily detect the configuration and location of IDS by analyzing these messages in the network. It is possible therefore to guard the IDS by encoding its messages or to create a separate network for management as shown in the

ISSN 2348 – 7968

diagram. The advantage of having a separate network between IDS Manager and IDS Sensors is not only to provide security but also to ensure "out of band" communication, meaning no bandwidth of the existing network is utilized for its communication. It is generally recommended to use IDS sensors inside and outside the firewall or between each firewall in a multi-layered environment and host based IDS on all critical or key hosts. IDS Management Module and its sensors communicate via zero bandwidth LAN segment in a transparent or stealth operation mode. This kind arrangement enables the IDS to have complete view of the organizational network and can even detect the failed attempts of attacks while reducing the chances of being compromised.

2. Related Work:

Network intrusion detection systems like snort [3] or Bro [11] typically use signature based detection, matching patterns in network traffic to the patterns of known attacks. This works well, but has the obvious disadvantage of being vulnerable to novel attacks. An alternative approach is anomaly detection, which models normal traffic and signals any deviation from this model as suspicious. The idea is based on work by Forrest et al. (1996), who found that most UNIX processes make highly predictable sequences of system calls in normal use. Network anomaly detectors look for unusual traffic rather than unusual system calls. ADAM (Audit Data and Mining) [12] is an anomaly detector trained on both attack-free traffic and traffic with labelled attacks. It monitors port numbers, IP addresses and subnets, and TCP state. ADAM uses a naive Bayes classifier which means that the probability that a packet belongs to some class (normal, known attack, or unknown) depends on the a-priori probability of the class, and the combined probabilities of a large collection of rules under the assumption that they are independent. In the IDES/NIDES systems [9], [10], a statistical based anomaly detection technique is used to represent the expected normal behavior of a subject and variance due to noises. The statistical-based anomaly detection technique overcomes the problems with rule-based anomaly detection technique in handling noises and variances. However, the statistical technique in IDES/NIDES is a univariate technique that is applied to only one behavior measure, where as many intrusions involve multiple subjects and multiple actions having impact on multiple behavior measures. Hence, a multivariate anomaly detection technique is needed for intrusion detection. Matthew V. Mahoney and Philip K. Chan developed "Packet Header Anomaly detection for identifying Hostile Network (PHAD)" [16],[17] that learns the normal ranges of values for each packet header field at the data link (Ethernet), network (IP), and transport/control layers (TCP, UDP, ICMP). PHAD detects some of the attacks in the DARPA data set that involve exploits at the transport layer and below. The paper, "Detecting Novel Network Intrusions Using Bayes Estimators" [18] authored by Daniel Barbara and et al suggests a method called pseudo-Bayes estimators as a means to estimate the prior and posterior probabilities of new attacks.

Data mining based intrusion detection techniques can be classified into two categories: misuse detection and anomaly detection. In anomaly detection technique, models are built on normal behavior and any deviation from normal behavior is identified as intrusion [2].Anomaly detection tries to determine whether deviation from normal usage pattern can be flagged as intrusion. It establishes normal usage patterns using statistical measures on system audit data and

ISSN 2348 – 7968

network data. Though new kinds of intrusions are detected, this benefit is paralyzed by high number of false alarms. More over improper/ insufficient training to anomaly module results in showing the genuine changes in the network traffic pattern as suspicious activities only to raise the number of false positives and false negatives..In misuse detection technique, each instance in a dataset is labelled either as 'normal' or 'intrusion' and learning algorithm is trained over labelled data to build model. Whenever a new type of attack is discovered, learning algorithm can be retrained with new dataset that includes labelled instances of new attack. In this way, models of misuse detection are created automatically and can be more precise than manually created signatures. The rest of this paper discuss about the details of the proposed system.

3. Methodology:

3.1 System Architecture:

The following diagram shows the overall process of the proposed system. A classification technique, relevance vector machine (RVM) model combining kernel principal component analysis (KPCA) with particle swarm optimization (PSO) is proposed for intrusion detection in the network based environment. In the proposed model, a multi-layer RVM classifier is adopted to estimate whether the action is an attack, KPCA is used as a preprocessor of RVM to reduce the dimension of feature vectors and to shorten the training time. In order to improve the performance of RVM, an improved kernel function is proposed. PSO is employed to optimize the kernel parameters. The sparseness property of RVM allows automatic selection of the right kernel at every location by pruning all digressive kernels.

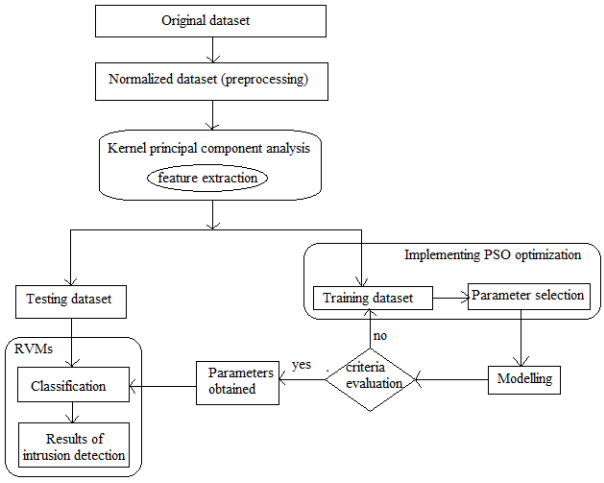

Fig: 1 The procedure of the Proposed RVM model for intrusion detection.

3.2 KERNEL PRINCIPAL COMPONENT ANALYSIS

Principal component analysis (PCA) is a common method applied to dimensionality reduction and feature extraction. PCA method can only extract the linear structure information in the data set, however, it cannot extract this nonlinear structure information. KPCA is an improved PCA, which extracts the principal components by adopting a nonlinear kernel method. A key insight behind KPCA is to transform the input data into a high dimensional feature space in which PCA is carried out.PCA begins by computing the covariance matrix of the *m* x *n*matrix X

$$C = \frac{1}{m}\sum_{i=1}^{m} \mathbf{x}_i \mathbf{x}_i^{\mathsf{T}}.$$

It then projects the data onto the first *k* eigenvectors of that matrix. By comparison, KPCA begins by computing the covariance matrix of the data after being transformed into a higher-dimensional space,

ISSN 2348 – 7968

$$C = \frac{1}{m}\sum_{i=1}^{m}\Phi(\mathbf{x}_i)\Phi(\mathbf{x}_i)^{\mathsf{T}}.$$

It then projects the transformed data onto the first *k* eigenvectors of that matrix.

3.3 RVM BASED ON PSO

Relevance vector machines (RVM) have recently attracted much interest in the research community because they provide a number of advantages. They are based on a Bayesian formulation of a linear model with an appropriate prior that results in a sparse representation. As a consequence, they can generalize well and provide inferences at low computational cost.

Suppose there are N particles, and the updating and changing equation of the particles' speeds and positions are as follows:

$$v_i = v_{i-1} + C_1 . r_1 . (pbest - x_{i-1}) + C_2 . r_2 . (gbest - x_{i-1}) \quad (1)$$

$$x_i = x_{i-1} + v_i \quad (2)$$

Here x_i and v_i (i = 1, 2, . . . N) are respectively the position and speed of the particles, C_1 and C_2 are random variable quantities [1] in the scope of [0,1]; *pbest* and *gbest* respectively stand for right now the best position of the particles and the best detecting position of all the particles.

Fix V_{max} the maximum speed, and then the speed of the particle can be changed as follows:

$$V_i = V_{max}, \quad \text{if } V_i > V_{max}$$
$$V_i = -V_{max}, \quad \text{if } V_i \le V_{max} \quad (3)$$

Adopting PSO calculation can get the best RVM parameters at the best accuracy, and the steps are as follows:

Step1: Randomly beginning the particle group; the speed and position of every particle are begun at random, calculate the adaptation function of every particle, and the beginning valueof$pbest_i$, and the beginning value of the particle are the same, and then the beginning of *gbest* is the particle with best function value.

The definition of the adaptation function is:

$$F_{fitness} = \sum_{i=1}^{N}\left[\frac{y-y_i}{N}\right]^2 \quad (4)$$

In this equation, yis the examination value of the sample, $y_i$is the prediction value of the sample, and N is the number of the samples.

Step2: Change and update the particles; the speed of each particle and the update of the particle respectively follow Equation(1) and Equation (2) , and calculate the adaptation scope of each particle to find the best position .

Step3: Repeat step2 till reaching the maximum changing times or the changing astringent accuracy allowed.

4. Results:

In this section, we selected samples from the subset of KDD to form the training and testing set. There are some performance indicators for the intrusion detection system as follows: TP, FP, TN, FN, where TP represents that the normal behavior is correctly forecasted, FP indicates that the abnormal behavior is judged as normal, FN denotes that the normal behavior is wrongly thought as abnormal, and TN represents the abnormal behavior is correctly detected. The table 4.1 shows the accuracy, training and

ISSN 2348 – 7968

testing time needed for the proposed intrusion detection system.

	KDD_DATASET	Accuracy	train(seconds)	test(seconds)
RVM-PSO	53	94	66	0.048
Out of	269 (2412)			

Table 4.1. Performance of proposed system

TRUE POSITIVE

The True Positive (TP) is defined as number of correctly detected iris image to the total number of images. The TP formula is defined in eq.(1).

$$TP = \frac{Number\ of\ correctly\ Detected\ Iris\ \text{Im}ages}{Total\ No.of\ \text{Im}ages} \text{ x } 100 \quad \text{.....(1)}$$

TRUE NEGATIVE

The True Negative (TN) is defined as number of falsely detected iris images to the total number of images. The TN formula is defined in eq.(2).

$$TN = \frac{NumberOfFalselyDetectedIris\ \text{Im}ages}{TotalNoof\ \text{Im}ages} \text{ x } 100 \quad \text{.....(2)}$$

FALSE POSITIVE

The False Positive (FP) is defined as number of correctly detected non-iris images to the total number of images. The FP formula is defined in eq.(5.3).

$$FP = \frac{NumberOfCorrectlyDetectedNonIris\ \text{Im}ages}{TotalNoof\ \text{Im}ages} \text{ x } 100 \quad \text{.....(3)}$$

FALSE NEGATIVE

The False Negative (FN) is defined as number of falsely detected non-iris images to the total number of images. The FN formula is defined in eq.(4).

$$FN = \frac{NumberOfFalselyDetectedNonIris\ \text{Im}ages}{TotalNoof\ \text{Im}ages} \text{ x } 100 \quad \text{.....(4)}$$

Accuracy:

Accuracy is the measurement system, which measure the degree of closeness of measurement between the original value and the extracted value.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

Where, TP – True Positive (equivalent with hit)

FN – False Negative (equivalent with miss)

TN – True Negative (equivalent with correct rejection)

FP – False Positive (equivalent with false alarm

Table 4.2 shows the result of different methods. In the table, we can see that in this paper we use Particle swarm optimization to optimize the parameter, Kernel principal component analysis method is used to reduce dimension. Compared to WPSO-SVM, it has better accuracy, because of the kernel principal component analysis method can extract nonlinear feature. Compared to KPCA-PSO-SVM, we also got a good result, because RVM uses Bayesian inference to obtain parsimonious solutions for probabilistic classification. From the table we also can see the speed of KPCA-PSO-SVM is slower than the proposed KPCA-PSO-RVM.

ISSN 2348 – 7968

Method	Accuracy (%)	Testing time (s)
KPCA-PSO-RVM	94.0	0.048
KPCA-PSO-SVM	93.604	0.520540
WPSO-SVM	93.0693	0.04487
PSO-SVM	92.4752	0.035042

Table 4.2. Comparison of results

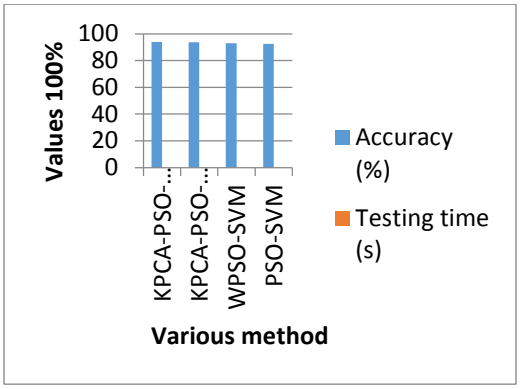

Fig: 2. Performance Graph

5. CONCLUSION

Network Intrusion Detection System has a major role to play in safeguarding the network resources against various kinds of attacks. With the advent of new vulnerabilities and sophistications in the nature of attacks, new techniques for intrusion detection have evolved. The main objectives of the research being increasing the detection accuracy while keeping the false positive rate low. In the present framework of project, discussed the design and development of "A novel Intrusion Detection System based on KPCA and RVM with PSO model" which is built using PSO that helps in selecting parameters of the RVM in the intrusion detection. We combine KPCA with PSO-RVM and apply it to intrusion detection system. The theory and results confirmed that our method do better in generalization and accuracy.Presently, the work caters only to identify and classify the events into normal and attack classes. It can be extended to detect and classify the attacks into multiple attack classesfor future enhancement.

REFERENCES

[1]. S.Suresh, P.B.Sujit, A.K.Rao. Particle swarm optimization approach for multi objective composite box-beam design [J]. Composite Structures, 2007, 81(4): 598-605

[2]. J. P. Anderson, "Computer Security Threat Monitoring and Surveillance", Technical Report April 1980, http://csrc.nist.gov/publications/history/ande80.pdf

[3]. Martin Roesch : "Snort Documents", http://www.snort.org/docs/

[4]. Net Optics, Inc. "White Paper: Deploying Network Taps with Intrusion Detection Systems", http://www.netoptics.com/products/downloads.asp?PageID=150&Section=res

[5]. K. Kendall, "A Database of Computer Attacks for the Evaluation of Intrusion Detection Systems," Massachusetts Institute of Technology Master's Thesis, 1998.

[6]. Basic Analysis and Security Engine project, http://base.secureideas.net/

[7]. Chang-Tien Lu,Arnold P.Boedihardjo,Prajwal Manalwar, "Exploiting efficient data mining techniques to enhance Intrusion Detection Systems," 0-7803-9093-8/05/$20.00 2005 IEEE, pp. 512-517.

[8]. Brugger S. T, "Data mining methods for network intrusion detection," Technique Report, UC Davis, 2004.

[9].Javitz HS, Valdes A. "The NIDES statistical component description of justification" Technical Report A010, SRI International, Menlo Park, CA, March 1994. http://www.cs.ucdavis.edu/~wu/ecs236/papers/hw2_NIDES-STA-description.pdf

[10]. Javitz HS, Valdes A. "The SRI statistical anomaly detector", Proceedings of the 1991 IEEE Symposium on Research in Security and Privacy, May 1991

[11].V. Paxson, "Bro: A System for Detecting Network Intruders in Real-Time", Computer Networks, 1999, http://bro-ids.org/publications.html

[12]. D. Barbar´a and S. Jajodia and N. Wu and B. Speegle, "The ADAM project", http://www.isse.gmu.edu/dbarbara/adam.html

[13]. Tipping M E. Sparse Bayesian Learning and the Relevance Vector Machine [J].Journal of machine learning research.2001, 1(3):211-244

[14]. S. Stofo et al. The Third International Knowledge Discovery and Data Mining Tools Competition, The University of California, 2002 [Online]. Available: http://kdd.ics.uci.edu/databases/kddCup99/kddCup99.h-tml.

[15]. S. Mukkamala, A. H. Sung, and A. Abraham, "Intrusion detection using an ensemble of intelligent paradigms," Netw. Comput. Appl., vol. 28, no. 2, pp. 167–182, Apr. 2005.

[16].M. Mahoney and P. Chan, "PHAD: Packet header anomaly detection for identifying hostile network traffic", Technical report, Florida Tech., technical report CS-2001-4, April 2001, http://citeseer.ist.psu.edu/mahoney01phad.html

[17]. Mahoney M. and P. Chan, "Learning models of network traffic for detecting novel attacks", Technical report, Florida Tech 2002, http://cs.fit.edu/~mmahoney/paper5.pdf

[18]. D. Barbara, N. Wu and S. Jajodia, "Detecting Novel Network Intrusions using Bayes Estimators", Proceedings of the 1st SIAM International Conference on Data Mining, 2001, http://www.cs.ubc.ca/local/reading/proceedings/siam_datamining2001/pdf/sdm0129.pdf

[19]. W. Wang, R. Battiti, Identifying intrusions in computer networks with principal component analysis, in: Proceedings of the First International Conference on Availability Reliability and Security (ARES'06), 2006, p.270-279.

[20]. L. Khan, M. Awad, B. Thuraisingham, A new intrusion detection system using support vector machines and hierarchical clustering, int. J. Very Data Bases 16 (2007) 507-521.

[21]. C.F. Tsai, Y.F. Hsu, C.Y. Lin, W.Y. Lin, intrusion detection by machine learning; a review, Expert Syst. Appl.36 (2009) 11994-12000.

[22]. S.X. Wu, W. Banzhaf, Use of computational intelligence in intrusion detection systems: a review, Appl. Soft comput. 10 (1) (2010) 1-35.

[23]. Li Guo-dong, 1,*Xia Ke-wen, 1Zhao Qian-qian, 1Bai Jian-chuan, "A Study in Intrusion Detection Based on Relevance Vector Machine Optimized" International Journal of Advancements in Computing Technology(IJACT) Volume5,Number8,April 2013.

ISSN 2348 – 7968

[24]. Pritam Sapate 1 and Shital A. Raut 2 "survey on classification techniques for intrusion detection" Dhinaharan Nagamalai et al. (Eds): ACITY, WiMoN, CSIA, AIAA, DPPR, NECO, InWeS – 2014.

[25]. Lei Li a, Kongyuan Liu b, An Intrusion Detection Method Based on WPSO-SVM and KPCA. Journal of Information & Computational Science 11:5 (2014) 1403–1410.