

AN APPROACH FOR IDS BY COMBINING SVM AND ANT COLONY ALGORITHM

Subaira.A.S¹, Anitha.P², Anjusree.S³, Vinothini.C⁴

¹PG Scholar, CSE, Dr.N.G.P.Institute of Technology, Tamilnadu, India

²Assistant Professor, CSE, Dr.N.G.P.Institute of Technology, Tamilnadu, India

³PG Scholar, CSE, Dr.N.G.P.Institute of Technology, Tamilnadu, India

⁴Assistant Professor, CSE, Dr.N.G.P.Institute of Technology, Tamilnadu, India

Abstract

This piece of work researches the intrusion detection problem of the network sanctuary; the primary task is to classify network behavior as normal or abnormal while reducing misclassification. In this paper, two efficient data mining algorithms are combined together to detect the network intrusion. Combining SVM and Ant colony (CSVAC) used for well-organized data classification, this technique takes the advantage of both the algorithm while avoiding their weaknesses. This algorithm is implemented and evaluated using standard benchmark KDDCUP99 data set. Experimental results drastically well produce superior results than the other algorithm in terms of accuracy rate and run time efficiency, and this algorithm able to detect the new types of attacks

Keywords: Intrusion Detection; Support Vector Machine; Ant colony; Combined Support vector with ant colony

-----***-----

1. INTRODUCTION

In today's information system security as remained one hard line area for both computers and networks. Although different tackles and methods have been proposed to handle the security problem, very few are adequate and efficient enough for real applications. The Intrusion Detection system is used to protect the system from various types of computer viruses and attacks. To protect the computers and networks from various cyber-attacks and viruses the Intrusion Detection Systems (IDS) are designed. An IDS is a mechanism that monitors network or system actions for malicious activities and produces reports to a management station [1]. IDSs build efficient clustering and classification models to distinguish the normal behaviour from abnormal behaviour using data mining techniques. The two important aspects of the research of intrusion detection are, first the user program activities are monitored by a computer system via system auditing technique, second normal and intrusion activities must have dissimilar behaviours [2].

Several existing IDS cannot efficiently deal with new types of viruses, attacks for varying computing environments. So the installed IDSs for ever and a day need to be restructured. Since it is energy and time consuming job. IDS discover the behaviors of networks routinely by analyzing the data trails of their activities by data mining approaches. The two main advantages of using a data mining approach toward IDSs are first it can be used to automatically generate the detection models for IDS so that new attacks can be detected automatically. Second it can be used to build IDSs for a wide variety of computing environments.

2. RELATED WORK

In this Paper [11] Machine learning based data classification techniques has been used for intrusion detection. It included established classification methods such as K-Nearest Neighbor (K-NN), Support Vector Machines (SVM), Decision Trees (DT), Bayesian Method, Self-Organized Map (SOM), Neural Networks (NN), Generic Algorithms (GA), and Fuzzy Logic techniques to advance the performance of the classifier [11]. As declared in [12] a computational intelligence was used for intrusion detection System. It offered the details of the classification algorithms that cover core procedure of CI with artificial neural networks, fuzzy systems, swarm intelligence, and soft computing. Another evaluation Stated in [13] based on integrating On Line Analytical Processing (OLAP) tools and data mining techniques for identifying Intrusion Detection System. This paper shows that the relationship of the two fields produces a good resolution to deal with defects of IDSs such as low detection accuracy and high false alarm rate. In this paper [14] Naïve Bayes algorithm and Support Vector Machine algorithm outcome have been used to compare the detection rates and false alarm rates of intrusion detection system. Also the Naïve Bayes method with Ant colony Optimization technique also proposed to improve the rates of detection.

This work [15] is based on self-organized ANT colony Intrusion Detection System (ANTIDS) used to detect intrusions in a network infrastructure. Then the performance is compared between Decision Trees, Support Vector Machines and Genetic Programming for efficient intrusion detection

systems. In this paper [16] SI method have been used for constructing IDS main contribution of this work is a detailed comparison of several SI-based IDS in terms of efficiency. This presents a clear idea of which solution is more suitable for each particular case. In this paper [18] intelligent learning approach using Ant Colony Optimization (ACO) is used for identifying intrusion in the distributed network. The Experimental outcome is used to detect the hidden intrusion attacks with high detection rate and low false alarm rate. This paper [3] BIRCH is based on handling an enormous data set in the computer memory and train SVM on the tree's nodes. After building the hierarchical tree the training process of SVM starts and it causes expensive computational risk, particularly when the data is not uniformly distributed or cannot fit in the computer memory. The main purpose of the clustering algorithm is to reduce the expensive disk access cost.

In this work [16] hybridization of evolutionary fuzzy systems and ant colony optimization which is used for Intrusion Detection. In this paper [6] DT and support vector machine algorithm applied to built two classifiers for comparison by employing a sampling method of several different normal data ratios. In this paper [28], RST (Rough Set Theory) and SVM is used to detect intrusions. The preprocessing and feature selection is controlled by RST. The selected features were sent to SVM model to learn and test respectively. This process is effective to decrease the space density of data. The experiments will evaluate the results with different methods and explain RST and SVM schema might improve the false positive rate and accuracy. Much study has been done on neural network for intrusion detection [23], [24], [25]. Neural network is often proposed as the statistical analysis component of an anomaly detection system. In [26] A MODEL for recognizing abnormal behavior of a user on a computer system using neural network is presented. A neural network based IDS will first use the training data including sequence of normal activities to train the neural network, and then run the neural network to predict normal activity. The neural networks learned can be much more complicated than the rules learned by decision tree methods.

3. COIMBINED SUPPORT VECTOR WITH ANT COLONY APPROACH FOR INTRUSION DETECTION

In this section the techniques like Support Vector Machine, Ant Colony and CSVAC algorithm has been described in detail for intrusion detection. It shows how classification and clustering can be used for IDS.

3.1 KDDCUP99 Data Set

The performance was evaluated by the standard bench mark data set KDDCUP99 provided by the MIT Lincoln laboratory.

The data set contains different types of intrusions present in networking environment. Mainly TCP/IP data combined with several attacks. Each TCP/IP connection contains features like duration, type of protocol, Src_bytes, dst_bytes, dst_hst_service_count; etc., each data item is labeled as either normal or an attack with any of following four categories (DoS, U2R, R2L, Probe).The classification of dataset was present in Table 1.

The four major attacks are (i)DoS(Denial of Service): A denial-of-service attack or distributed denial-of-service attack is an effort to make a computer source out of stock to its legitimate users[24] .It Make the system to slows down or shut down so it interrupt the service and rebuff the indented authorized user. Due to this attack high network traffic occurs [15] (ii) User to Root Attack (U2R): In this attack, the attacker starts at client level like snatching the password, dictionary attack and at last attacker achieves the root to access the system (iii) Remote to User Attack (R2U) In this attack, The attacker can produce vulnerability over a network and have the ability to send a packet over a network which does not have an account on that machine. (iv)Probe: In this type of attack, an attacker examines a network to gather information or discover well-known vulnerabilities

Table 1: Classification of Data set

Category	Attack type
Normal	Normal(39298)
DOS	Smurf(11258), Neptune(11954), Back(2001), Teardrop(99), Pod(20), Land(17)
U2R	Buffer overflow(3), Root kit(10), load module(7), Perl(2)
R2L	Warezcilent(1020), Guess_passwd(53), Warezmater(20),Imap(9), ftp_write(8), Multihop(5), Phf(1), Spy(2)
Probe	Satan(20), Ipsweep(658), Portsweep(40), Nmap(130)

3.2 SVM Classification

SVM is a learning method for the Classification and Regression analysis of both linear and nonlinear data. It uses a hypothesis space of linear functions and maps input feature vectors into a higher dimensional space all the way through some nonlinear mapping [2].SVM constructs a hyper plane or set of hyper planes only the good separation is achieved by the hyper plane. The hyper plane searching process in SVM is achieved by the leading margin [7] [13]. The related margin gives the major separation between classes. While training an SVM it creates a quadratic optimization problem [4].

In SVM the classifier is created by linear separating hyper plane but all the linear separation cannot be solved in the original input space. SVM uses a function called kernel to solve this problem. The Kernel transforms linear problem into nonlinear one by mapping into feature spaces. Intrusion detection system has two phases: training and testing. SVMs can learn a larger set of patterns and be able to provide better classification, because the categorization difficulty does not depend on the dimensionality of the feature space. SVMs also have the ability to update the training patterns dynamically whenever there is a new pattern during classification [11]. Linear classification adds the data point belong to either class A or class B. Training data point x_i can be labelled by y_i based on (1)

$$y_i = \begin{cases} -1 & x_i \in \text{class A,} \\ 1 & x_i \in \text{class B.} \end{cases} \quad (1)$$

For SVM training, the data point can be taken from network connecting record that describes several features. The conventional SVM algorithm is operated above the entire training data set. The dimension of the matrix for computing kernel functions can be determined from the number of training data points. This kernel function influences the time of solving the QP problems. SVM have one attractive feature that the data points do not lay on the margin do not be taken for computation. So, the number of training data points can be condensed without trailing accuracy.

3.3 Clustering

The ant colony algorithm can suggest very interesting result for intrusion detection problem. In natural the ants have the intellectual character. This algorithm describes the Ant system manner based on such nature of ants and it produced great result. After performing ant clustering phase the patent objects are stored in the SVM training data file for further process. The heuristic features of this algorithm are robustness, distributed and parallel computing feature and positive feedback characteristic. In this process the similarity of two objects is calculated. The distance function of two similar object is $d(O_i, O_j)$, the Euclidean distance, for example. The decreasing function of two similarity object O_i and O_j can be described by a as follows

$$S(O_i, O_j) = 1 - \frac{d(O_i, O_j)}{\beta} \quad (2)$$

Where β is a positive coefficient. The ant colony algorithm classify the objects into different classes the first one is normal and the remaining classes are different kind of intrusions. The intrusion detection classifier can be built based on anomaly and misuse detection pattern. The rule of Ant clustering based intrusion detection is described below. The swarm similarity coefficient β is adjusted, so the ant clustering algorithm clusters the training data set into s clusters $\{c_1, c_2, \dots, c_s\}$, c_i

$= \{O_{ij}\}$, where $i = 1, 2, \dots, s, j = 1, 2, \dots, n_i$, O_{ij} are the objects that belong to cluster c_i and n_i is the number of objects of c_i [base]. The midpoint of each cluster c_i is calculated as $C_i = \text{average} \{O_{ij}\}$ for all $O_{ij} \in c_i, i = 1, 2, \dots, s$. Consider that there are n classes in the training data set. The different clusters might belong to the same class of data. The Average of cluster center can be calculated by

$$A_k = \text{average} \{C_i\}, \text{ for all } c_i \in \{\text{Class } k\}, k = 1, 2, \dots, l.$$

3.4 Combining SVM with Ant Colony

In this phase the new machine learning algorithm has been introduced namely Combined Support Vector and Ant Colony (CSVAC). It is based on a mixture of the customized version of the two algorithms discussed above (SVM and AC). The interactive algorithm SVM and AC are taken multiple times at this phase for intrusion detection. At first SVM finds the support vectors and then generate hyper plane that is used to separate normal and as well as for each class of abnormal data while an AC is used to discover data added to the SVM training set. At last ant colony create models for normal data and abnormal data.

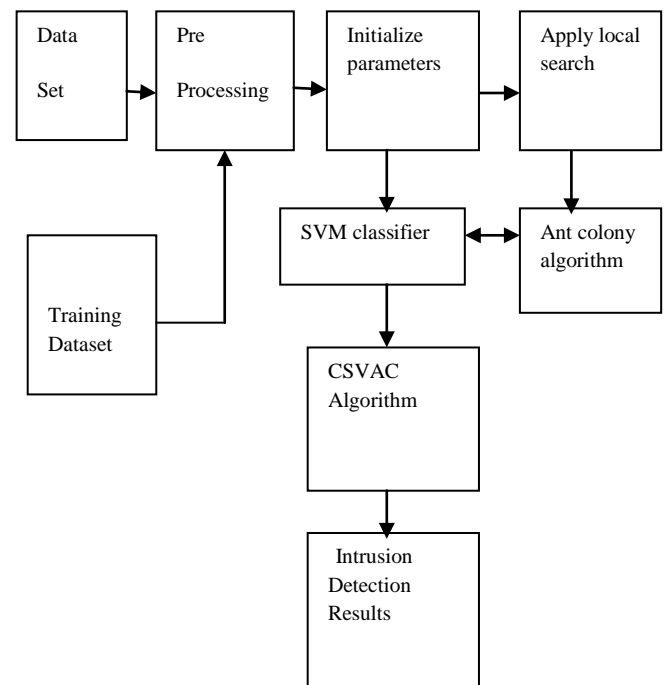


Fig 1: Block diagram for IDS using CSVAC

Algorithm 1: SVM with Ant clustering

Input: Training set with each data point labeled as normal or abnormal (class labels).

Output: A classifier.

1 **begin**

2 Randomly choose data points from each class.

3 Generate a SVM classifier.

- 4 **while** n number of data points added to training set do
- 5 locate support vectors among the chosen points;
- 6 Apply Ant clustering around the support vectors;
- 7 Add the points in the clusters to the training set;
- 8 Retrain the SVM classifier using the updated training set;
- 9 **end**
- 10 **end**

The active learning SVM process is introduced at this stage; it repeatedly performs the SVM training based on different training data sets. The main issue of active learning SVM is how to pick the training data set for each training step. The separating hyperplane is used to handle this issue; hyperplanes are generated by support vectors. The hyperplane can be tailored gradually by adding points between the margins only after each SVM training phase. The main objective of active learning SVM is to find the support vectors between the entire training data points. Hence it produces more efficient selection strategy of data points.

Algorithm 2: Training in CSVAC

Input: A training data set.

Input: TI – number of training iterations.

Input: DT – detection rate threshold.

Output: SVM and AC Classifiers.

- 1 **begin**
- 2 Normalize the data;
- 3 Let DR is the detection rate, initially 0;
- 4 **while** DR < DT do
- 5 **for** k = 1, . . . , TI do
- 6 SVM training phase;
- 7 Ant clustering phase;
- 8 **end**
- 9 Construct classifiers;
- 10 Do testing to update DR;
- 11 **end**
- 12 **end**

The clustering technique is needed for data selection process in active learning SVM. Ant Clustering is a more preferable choice because the classification process can be taken place in real time also it does not require any retraining process when the new training data has been added. False negatives can be decreased only when the new data is confirmed as normal by both of the classifiers. When both the classifiers prove that the data item is abnormal, then the subclass of the abnormal data item, that is category of the intrusion can be determined by the ant clustering classifiers. If the classified results are not consistent by both the classifiers, then the data item is patent as amphibious. The amphibious can be described as new type of intrusion.

4. EXPERIMENTAL SETUP AND RESULT

4.1 Experimental Setup

It was written in C# .net and its performance was evaluated by standard benchmark KDDCUP 99 data set. The performance evaluation was based on training time and detection rate. To implement the program we used windows based computer with Pentium core processor 2.0GHZ, 250GB HDD, 2 GB of RAM.

4.2 Results

By applying CSVAC Algorithm on selected features set it finds the attacks accurately. From the exceeding implementation we have effectively produce some rules those categorize the declared attack connection. The training phase and testing phase are executed by two independent modules. The new algorithm processes the training and testing phases parallel in the new IDS. This is an important ability of IDSs that are intended for real-time detection.

Class	HitMean	HitVariance	Num_failed_log	Num_failed_log	Logged_inMea	Logged_inVar	Num_comprom	Num_comprom	Root_shellMea	Root_shellVar	Su_attempts
back	1.9600339601	0.038252422	0	0	1	0	0.9695204695	0.029553013	0	0	0
buffer_overflow	2.3333333333	1.3333333333	0	0	1	0	2.3333333333	0.3333333333	0.6666666666	0.3333333333	0
ftp_write	0.5	0.67140571	0	0	0.75	0.2142057142	0.125	0.125	0	0	0
guess_passwd	1.0568037735	0.020802438	1.0568037735	0.3236574746	0.018678245	0	0	0	0	0	0
imap	0	NaN	0	NaN	1	NaN	16	NaN	0	NaN	0
ipwweep	0	0	0	0	0.008704273	0.0082512692	0	0	0	0	0
land	0	NaN	0	NaN	0	NaN	0	NaN	0	NaN	0
loadmodule	2	NaN	0	NaN	1	NaN	1	NaN	1	NaN	0
multihop	0.6	0.8	0	0	0.4	0.3	0	0	0	0	0
neptune	0	0	0	0	0	0	0	0	0	0	0
nmmap	0	0	0	0	0	0	0	0	0	0	0
normal	0.0458166440	0.7862232795	0.000229192	0.0005632339	0.9195837831	0.0729471347	0.020102042	14.971323880	0.0003016880	0.0003016880	7.6339762837
perl	0	NaN	0	NaN	1	NaN	0	NaN	1	NaN	0
pgf	2	NaN	0	NaN	1	NaN	0	NaN	1	NaN	0
pod	0	0	0	0	0	0	0	0	0	0	0

Fig 2:SVM Classification Phase

F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
0	7	138	0	2	0	0	0	0	0	0	0
1	663	325	0	2	1	0	0	0	0	0	0
20	102	2063	0	2	1	0	0	0	0	0	0
0	9	36	0	2	1	0	0	0	0	0	0
0	245	2202	0	2	1	0	0	0	0	0	0
0	314	16151	0	2	1	0	0	0	0	0	0
0	320	18759	0	2	1	0	0	0	0	0	0
0	312	18759	0	2	1	0	0	0	0	0	0

Fig 3: Ant Clustering Phase

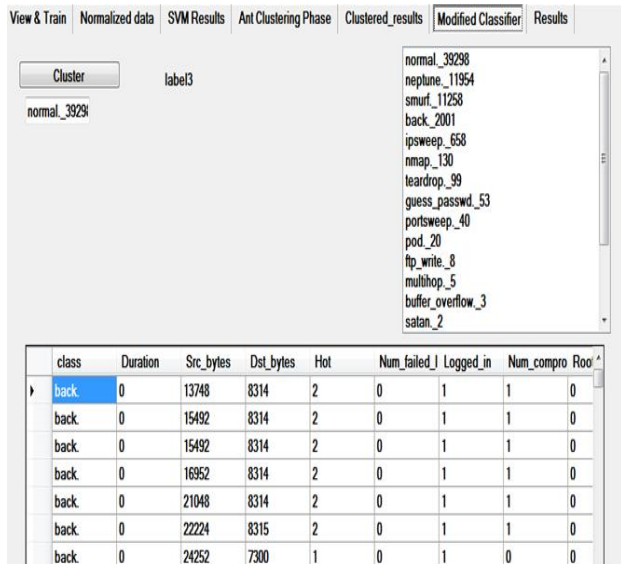


Fig 4:CSVAC phase

To compare the system performance under the three different modes (SVM only, CSOACN only, CSVAC), the classifiers generated by them should be tested separately by the same testing data set. As the general comparison results do not depend on the amount of testing data and the distribution of each data class, we chose the amount of testing data to be about 10 times the amount of training data. Let T1 be the testing data used for the comparison of the three algorithms. The class distributions of T1 are also shown in Table 2.

Table 2: Class allocation in network connection records

Class	KDD99 data set	Training set D	Testing set T1
Normal	39298	150	1200
DoS	25349	40	500
U2R	22	15	62
R2L	1118	60	1000
Probe	848	40	400
Total	66635	305	3162

The following three tables represent the confusion matrix of data set T₁.By using the CSVAC classifier that was constructed upon the data set D; all data in set T1 can be classified. Confusion matrices of data set T1 is given below. The testing results are shown in Table 3(c).The Training time, Detection rate, False positive and False negative between the above three algorithms has been shown in table 4.In this SVM have less training time compare with Ant Clustering where as Ant clustering has high detection rate compare with SVM.While combining these two algorithms it produce higher result.

Table 3.a: By SVM Classifier

Class	Normal	DoS	U2R	R2L	Probe
Normal	1028	0	9	0	0
DoS	0	436	0	0	0
U2R	0	0	47	0	0
R2L	1	0	424	486	0
Probe	0	0	0	0	107

Table 3.b: By AC Classifier

Class	Normal	DoS	U2R	R2L	Probe
Normal	927	15	3	42	13
DoS	0	493	0	0	7
U2R	0	0	39	13	0
R2L	45	312	12	628	3
Probe	3	257	20	2	218

Table 3.c: By CSVAC Classifier

Class	Normal	DoS	U2R	R2L	Probe
Normal	983	17	2	52	19
DoS	0	437	0	0	0
U2R	0	0	46	9	0
R2L	34	322	12	673	2
Probe	2	289	17	9	213

Table 4: Comparison of Performance Measure

Measure	SVM	AC	CSVAC
Training Time(s)	3.32	4.40	3.90
Detection Rate (%)	72.08	84.10	88.04
False Positive (%)	6.10	3.28	2.88
False Negative (%)	22.90	1.30	0.70

5. CONCLUSIONS

In this paper, a new machine learning based data classification algorithm Combined Support vector and ant colony has been established for the intrusion detection problem. In order to achieve superior performance two offered machine learning algorithms SVM and AC are combined to enhance accuracy rate and faster running time. The proposed Maximum Information Coefficient (MIC) method along with an enhanced CSVAC algorithm which is used to handle multiclass cases and to convert a nonlinear classification problem to a linear one is left for future work.

REFERENCES

- [1]. W. Lee, S.J. Stolfo, K.W. Mok, "A data mining framework for building intrusion detection models", in: Proceedings of IEEE Symposium on Security and Privacy, 1999, pp. 120–132.
- [2]. W. Feng, Q. Zhng, G. Hu, J Xiangji Huang," Mining network data for intrusion detection through combining SVMs with ant colony networks "Future Generation Computer Systems,2013.
- [3]. T. Zhang, R. Ramakrishnan, M. Livny, "BIRCH: an efficient data clustering method for very large databases", in: Proceedings of SIGMOD, ACM, 1996, pp. 103–114.
- [4]. L. Khan, M. Awad, B. Thuraisingham, "A new intrusion detection system using support vector machines and hierarchical clustering", The VLDB Journal 16(2007) 507–521.
- [5]. X. Xu, "Adaptive intrusion detection based on machine learning: feature extraction, classifier construction and sequential pattern prediction", Information Assurance and Security 4 (2006) 237–246.
- [6]. J.X. Huang, J. Miao, Ben He, "High performance query expansion using adaptive co –training", Information Processing & Management 49 (2) (2013) 441–453
- [7]. Y. Li u, X. Yu, J.X. Huang. A." An, Combining integrated sampling with SVM ensembles for learning from imbalanced datasets", Information Processing & Management 47 (4) (2011) 617–631.
- [8]. V. Vapnik, "The Nature of Statistical Learning Theory", Springer, 1999.
- [9]. P. Corsini, B. Lazzerini, F. Marcelloni, Combining "supervised and unsupervised learning for data clustering", Neural Computing & Applications 15 (3–4) (2006)289–297.
- [10]. C.-F. Tsai, Y.F. Hsu, C.Y. Lin, W.Y. Lin, "Intrusion detection by machine learning: a review", Expert Systems with Applications 36 (2009) 11994–12000.
- [11]. C.-F. Tsai, Y.-F. Hsu, C.-Y. Lin, W.-Y. Lin, "Intrusion detection by machine learning: a review", Expert Systems with Applications 36 (2009) 11994–12000.
- [12]. S.X. Wu, W. Banzhaf, "The use of computational intelligence in intrusion detection systems: a review", Applied Soft Computing 10 (2010) 1–35.
- [13]. H. Brahmi, I. Brahmi, and S.B. Yahia, "OMC-IDS: at the cross-roads of OLAP mining and intrusion detection", in: Advances in Knowledge Discovery and Data Mining, in: LNCS, vol. 7302, 2012, pp. 13–24.
- [14]. J. He, D. Long, C. Chen, "An improved ant-based classifier for intrusion detection", in: Proceedings of the 3rd International Conference on Natural Computation, Vol. 4, ICNC'07, IEEE Computer Society, 2007, pp. 819–823.
- [15]. V. Ramos, A. Abraham, "Antids: self organized ant-based clustering model for intrusion detection system", in: Proceedings of the 4th IEEE International Workshop on Soft Computing as Tran disciplinary Science and Technology,2005, pp. 977–986.
- [16]. C.-H. Tsang, S. Kwong, "Multi-agent intrusion detection system in industrial network using ant colony clustering approach and unsupervised feature extraction:, in: Proceedings of IEEE International Conference on Industrial Technology, IEEE Press, 2005, pp. 51–56.
- [17]. S. Freeman, J. Branch, "Host-based intrusion detection using user signatures", in: Proceedings of the Research Conference RPI., 2002.
- [18]. S. Janakiraman, V. Vasudevan, "ACO based distributed intrusion detection system", Journal of Digital Content Technology and its Applications 3 (1) (2009)66–72
- [19]. T.F. Lunt, "A survey of intrusion detection techniques", Computers and Security12 (4)(1993) 405–418.
- [20]. J. Ryan, M.-J. Lin, R. Miikkulainen,"Intrusion detection with neural networks", in: Proceedings of AAAI-97 Workshop on AI Approaches to Fraud Detection and Task Management, 1997, pp. 92–97
- [21]. H. Teng, K. Chen, S. Lu, "Security audit trail analysis using inductively generated predictive rules", in: Proceedings of the 6th Conference on Artificial Intelligence Applications, Vol. 1, 1990, pp. 24–29.
- [22]. D.E. Denning," An intrusion-detection model", IEEE Transactions on Software Engineering 13 (2) (1987) 222–232
- [23]. F. Monrose, A. Rubin, "Authentication via keystroke dynamics", in: Proceedings of the 4th ACM Conference on Computer and Communications Security, 1997
- [24]. Neri, F., "Comparing local search with respect to genetic evolution to detect intrusion in computer networks", In Proc. of the 2000 Congress on Evolutionary Computation CEC00, La Jolla, CA, pp. 238243. IEEE Press, 16–19 July, 2000.
- [25]. Neri, F. "Mining TCP/IP traffic for network intrusion detection", In R. L.de M'antaras and E. Plaza (Eds.), Proc. of Machine Learning: ECML\2000, 11th European Conference on Machine Learning, Volume 1810of Lecture Notes in Computer Science, Barcelona, Spain, pp. 313322.Springer, May 31- June 2, 2000.
- [26]. Dasgupta, D. and F. A. Gonzalez,"An intelligent decision support system for intrusion detection and response", In Proc. of International Workshop on Mathematical Methods, Models and Architectures for Computer Networks Security (MMM-ACNS), St.Petersburg. Springer-Verlag, 21-23 May, 2001
- [27]. Chittur, A., "Model generation for an intrusion detection system using genetic algorithms", High School Honors Thesis,

Ossining High School. In cooperation with Columbia Univ, 2001

[28]. Crosbie, M. and E. H. Spafford, "Active defense of a computer system agents", Technical Report CSD-TR- 95-008, Purdue Univ. West Lafayette, IN, 15 February 1995

[29]. UCI KDD Archive, KDD Cup 1999 data, 1999. <http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html>.

[30]. D. Duan, S. Chen, W. Yang, "Intrusion detection system based on support vector machine active learning", Computer Engineering 33 (1) (2007) 153–155.

[31]. V. Ramos, A. Abraham, "Evolving a stigmergic self-organized data-mining", in: Proceedings of the 4th International Conference on Intelligent Systems, Design, and Applications, 2004, pp. 725–730.