A Comprehensive Survey of Technologies for Building a Hybrid High Performance Intrusion Detection System

S.J. Sathish Aaron Joseph
Research Scholar and Head / Department of Computer Applications
J.J.College of Arts & Science (Autonomous)
Pudukkottai

R. Balasubramanian, Ph.D.
Professor and Research Guide, PG & Research Department of Computer Science
J.J.College of Arts & Science (Autonomous)
Pudukkottai

ABSTRACT
Intrusion detection plays a vital role in maintaining the stability of any network. The major requirements for any intrusion detection system are speed, accuracy and less memory.Though various intrusion detection methods are available, they excel at some points while lack in the others. This paper presents a comprehensive survey of the technologies that are used for detecting intrusions. It analyzes the pros and cons of each technology and the literature works that utilizes these technologies. Challenges faced by the current IDS and the requirements for IDS in the current network scenario are discussed in detail. A detailed study on the datasets that can be used for effective building of an IDS is discussed. The research framework is proposed and a discussion of the various technologies that can be used for improving the efficiency of the IDS is provided.

Keywords
Intrusion detection system; KDD CUP 99; SSENet; Evolutionary algorithms; Graph Database; Big Data

1. INTRODUCTION
Increase in the amount of data transfer in networked environments, especially the Internet has led to an increase in the potential threats. With the cost of processing getting decreased from time to time, adversaries are gaining more prominence and are exploiting the system vulnerabilities further. This has led to the development of mechanisms to counter the attacks, called the Intrusion Detection Systems (IDS). The major functionality of an IDS is to monitor and analyze traffic, identifying abnormal activities and assessing the severity of the situation and raising alarm. Figure 1 shows the architecture of a typical IDS.

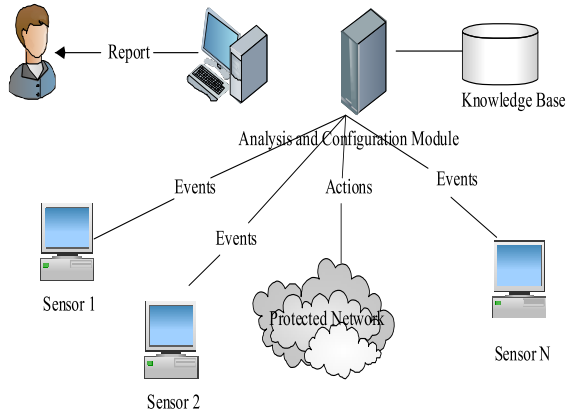

Fig 1: Architecture of typical IDS.

The major components of an IDS are the nodes/sensors on which the events take place. The events can correspond to normal activity or malicious activities. These events are recorded by the analysis & configuration module, which uses the knowledge base for categorizing the traffic as normal or anomalous. Reports are generated based on the analysis and is presented to the user for further analysis. The performance variables that play a vital role in determining the efficiency of the system are the detection rate (DR) and the false alarm rate (FAR). Since ID is basically a Classification problem, the ROC curve is used to determine the accuracy of the system.

1.1. Detection Methods
Intrusion detection methods are classified into two broad categories; anomaly based and misuse based (Figure 2)

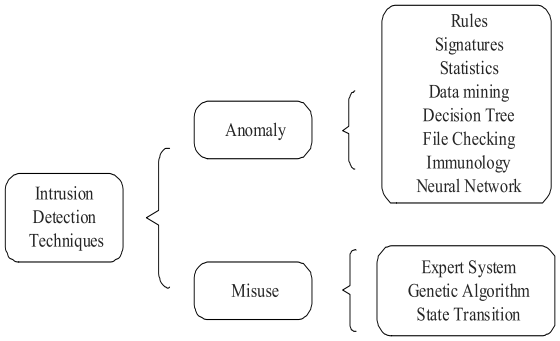

Fig2: Intrusion detection techniques

1.1.1. Misuse based IDS
Misuse based IDS uses already occurred attack patterns to identify attacks. Attack patterns that were previously encountered are coded as signatures and are maintained in the knowledge repository. Hence the misuse based IDS works on the already known attacks, while new attacks (patterns) cannot be diagnosed.

1.1.2. Anomaly based IDS
Anomaly based IDS uses normal instances as the base data to operate on. Any instance or behaviour deviating from this normal behaviour is termed anomalous and is categorized as an attack. This method does not use the previously available information, hence process speedup that can be achieved by this method is limited.

Apart from these defined techniques, various other techniques exists for intrusion detection. Target monitoring system deals with maintaining the file states and system status rather than monitoring for signatures or anomalies.Stealth probes IDS attempts to detect prolong attacks efficiently. Further,

intrusion detection techniques are categorized based on many different ways like statistics,neural network approach, data mining, genetic algorithm and computer immunology approach [31].

1.2. Detection Mechanisms

Intrusion detection mechanisms can be classified into two based on their area of operation. They are

• **Host-based (HIDS).**The HIDS resides on a host, monitoring traffics in and out of the particular host to identify malicious activities such as network events or system calls.

• **Network-based (NIDS).** NIDS, as the name suggests analyses the traffic of the entire network in which it is placed. The detection system resides on a single system andmonitors the network traffic analyzing the data for anomalies/patterns depending on the detection method in place. NIDS can be further classified as offline or online NIDS. Online or real time NIDS inspects the network packets to identify intrusions. Eg. Snort, Bro. Offline NIDS logs the details of the network traffic and analyses the log records (batch processing) to identify anomalies/patterns. NIDS is not suitable for detecting attacks that are launched on a specific host that are not launched through network.

The current trend in intrusion detection is to combine bothhost based and network based information to develop hybrid systems [30].

1.3. IDS VS IDPS

Intrusion Detection and Prevention System (IDPS) is an enhanced version of IDS, which performs both detection and also takes preventive measures based on the type of intrusions expected in the network. Figure 3 shows the architecture of an IDPS.

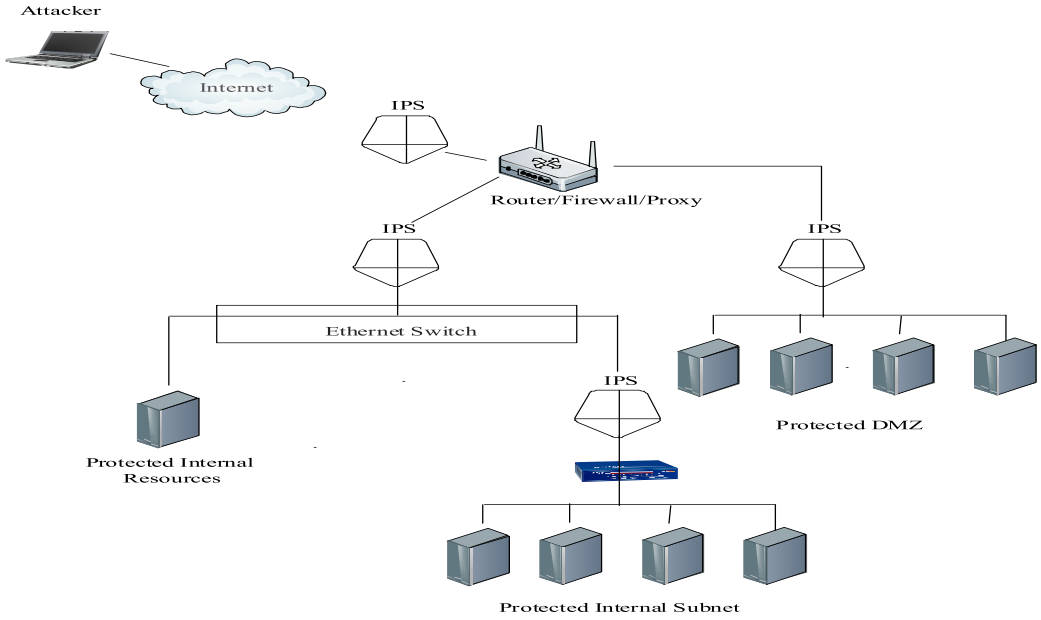

Fig 3: Network based Intrusion Prevention System

The internal resources (nodes)are protected by the IPS. The IPS connects directly to the internet/routers, safeguarding the internal network from outside attacks. It works on the basis of human immune system that protects the body from pathogens and viruses by altering its level of immunity [1]. Hence when considering networks requiring high availability, IPS plays a major role in the process of protecting the systems from outside attacks.

2. CHALLENGES FACING IDS

2.1. Performance Issues

Major issues corresponding to any Intrusion Detection System is its requirement for faster processing. Data from networks are small packets, but the number of packets transferred in the network per unit time is so huge that the job of the IDS becomes so complicated. Speed is a major requirement, since failure to detect intrusions in real time will lead to a missed opportunity, hence the intrusion detection system must be competent to produce results in a short time span, in other words an intrusion detection system must operate on real time to provide effective results. Accuracy is also one of the most important aspects in determining the performance and efficiency of the system. Accuracy here is measured in terms of the false positive rate, which is expected to be very low.

The load imposed by an IDS on the host system is also a major concern. Since the operations performed by IDS are basically pattern matching and hence is CPU intensive. Though improvements have been made on improving the efficiency of the string matching algorithms, it still remains to be imposing a heavy load on the system and the host system is virtually considered to be operating solitarily for the process of intrusion detection.

Another major issue is the I/O limitations imposed by the buffer. The network interface card is bounded by its buffer size, as only one host is employed with the IDS, it becomes mandatory for that system to analyze all the packets transferred in the network, which leads to buffer overflow [2].

The increasing flow of encrypted packets that is a major pitfall when utilizing signature based IDS. If the payload is encrypted, it becomes impossible to perform signature matching, which renders the existing signature repositories useless. Increasing new techniques and increased sophistication in attacks has lead to the need for more powerful IDS/IDPS.

2.2. Feature Selection /Feature Reduction

Network payload usually contains various fields corresponding to the packets being transferred. Not all features are useful for analysis. While some features might not contribute to the final result, some features might tend to bias the results in wrong directions. Hence it is not recommended to use the data as such. It becomes mandatory to analyze the data and clean it based on feature such that the final result set contains only the desirable data. While this becomes one of the mandatory components, it also imposes a level of overhead for the host in which the IDS is employed.

2.3. Attack Against IDS

Attacks on the nodes within the network and measures to counteract them are discussed in detail. But the point that is always missed out is that the node employing the IDS is also a part of the network, and it actually becomes the single point of failure. Though it is usually recommended that the IDS should be employed in a high performance and highly secure node, it always comes down to being a part of the network and no node is free/safe in a network. Compromising the node in which the IDS is employed will lead to a breakdown of the entire network. Methods to compensate for this circumstances is mandatory.

3. INTRUSION DETECTION TECHNIQUES: AN ANALYSIS

3.1. Evolutionary Algorithm based Intrusion Detection

Evolutionary/ Swarm Intelligence (SI) models seek inspiration from the behavior of a group of insects, birds or animals and their unique ability to solve problems as a swarm. The motivation behind these studies is that the natural phenomena can be directly mapped to the intrusion detection systems. A swarm of insects, even though primitive in knowledge, is able to perform even complex tasks due to the sharing of data and collaborative learning. This concept can be directly applied to the process of intrusion detection to perform effective analysis of the intrusions occurring in the network.

Swarm based techniques do not rely on the signatures/ patterns, hence they perform efficiently in detecting anomalous behavior. Further, IDS/IDPS techniques basically require processing of huge volumes of data and they also demand real time detection, which is possible in evolutionary methods.

Major techniques that utilize swarm intelligence methods are:

- ACO based approaches for Detecting the Origin of an Attack or for Induction of Classification Rules
- PSO Oriented Approaches using Neural Network, SVM and K-Means
- PSO for Induction of Classification Rules
- ACC Oriented hybrid approaches using SVM and SOM

It can also be observed from literature that Ant Colony based techniques are operated in isolation, while Particle Swarm techniques are always hybridized, by combining them with machine learning methods.

Table 1 shows the accuracy levels of swarm intelligence techniques compared with the accuracy levels of the winning entry of the KDD CUP 99 contest. It shows that the swarm intelligence techniques outperform in most of the attacks when compared with the algorithm developed in [5].

Table.1. Performance comparison of several SI-based IDS [32]

ML Type	**Winner KDD99**	**NN based techniques**				**SVM based techniques**			**Classification rules**			**None**			
SI Type		**PSO**				**PSO**			**PSO**	**ACO**		**ACC**			
	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]	[19]
Normal	94.5	N/A	96.88	N/A	N/A	N/A	N/A	N/A	N/A	98.5	96	99.64	98.5	98.8	99.1
Probe	83.3	88.86	92.20	N/A	N/A	86.48	N/A	N/A	N/A	82.5	86.25	98.29	86.9	87.5	97.18
DoS	97.1	92.57	97.74	N/A	N/A	88.48	N/A	N/A	N/A	98.5	98.83	99.98	97.5	97.3	99.35
U2R	13.2	91.14	52.86	N/A	N/A	85.52	N/A	N/A	N/A	76.3	72.8	64	27.2	30.7	63
R2L	8.4	94.29	8.30	N/A	N/A	84.53	N/A	N/A	N/A	89	33.45	99.47	11.0	12.6	97.79
DR	90.9	N/A	N/A	0.61	8.01	4.89	N/A	N/A	N/A	95.5	94.33	N/A	92.2	N/A	N/A
FAR	N/A	N/A	0.61	8.01	4.89	N/A	N/A	N/A	3.97	0.0018	N/A	N/A	1.5	N/A	N/A

3.2. Parallel Intrusion Detection using GPGPU

Computing systems have reached their limits in terms of memory wall and instruction level parallelism wall, due to the huge amount of processing involved. The process of intrusion detection, requires a huge amount of processing, which can be made efficient by parallelizing the detection process. Since the application involves same process (intrusion detection) to be carried out in different data, our application of intrusion detection is data parallelizable. Data parallelization is inbuilt in the multi core and many core processors, and can be performed with ease. Hence parallelization of the process of intrusion detection becomes possible. Though CPUs of the current generation contain many cores, the cost reduction in GPUs and the sheer number of processors (minimum 192 cores), easily outwits the possibility of CPU based processing and encourages GPU based processing [3]. The following are the theoretical advantages obtained by using GPUs for intrusion detection [4]:

- Provides better performance, memory bandwidth and parallelization capabilities;
- High scalability;
- Lower cost, hence economical;
- No real change required in the network structure.

Table.2. Comparison of IDS using GPU

Work	Year	Based on	Improving Method	Experimental Results
NIDS Pixel-Snort	2006	Snort	Off-load packet processing in detection engine to GPU	Decreases CPU load by 50%
Multiple- pattern Matching Algorithm	2008	Wu-Manber algorithm	Transformed into GPU as multi-pattern Matching algorithm	Twice computational performance
NIDS Gnort	2009	Snort	Off- load packet processing in detection engine to GPU	2,3 Gbit/s network traffic bandwidth
Framework for network traffic analysis	2010	CUDA	New framework for programming network traffic analysis	----
Accelerating the LOF algorithm	2010	Local outlier factor Algorithm	GPU implementation of the k-nearest neighbor algorithm to accelerate LOF classification	100x speedup than multi core CPU
Multi-Parallel IDS architecture	2011	Snort	Multi-parallel architecture	5,2 Gbit/s: network traffic 70 Gbit : pattern matching
NIDS Suricata	2011	Snort	Experimental CPU and CUDA technology support	10 Gbit/s network traffic bandwidth
Efficient Packet Pattern Matching	2012	Hierarchical hash table architecture on GPU	Balances the work load among the thread	10x than Aho-Corasick algorithm
Parullelizing NIDS	2012	Snort	Aho-Corasick algorithm on GPU	4x speedup than CPU implementation

Table 2 shows a comparison between GPU based techniques and their critical analysis. It shows that every method has its own focus and is not based on general improvements. Hence it can be understood that a tradeoff always occurs in terms of speed or memory or bandwidth consumption when performing GPU based intrusion detection.

3.3. Machine Learning based Intrusion Detection

Machine learning techniques in general refers to the construction of algorithms that performs predictions by learning from the data. Classification techniques are mostly supervised or semi-supervised, while Clustering is an example of unsupervised learning [20]. Hence machine learning techniques are capable of both anomaly and misuse detection [25]. Machine learning techniques, though being powerful, are

used in combination with other techniques, as discussed in section 3.1.

3.4. Graph based Intrusion Detection

Graph database is a structure that uses nodes and edges to represent data. The data represented as log records comprise of various related structures, and hence can be represented as a graph. The major advantage of a graph based analysis system is its ability to provide partition free tolerance. Since the data corresponding to intrusion detection can be very large (scalable with respect to the network), it becomes an added advantage. Graph databases are usually scalable and they tend to be fast, as the entire graph is used for processing.

The existing graph based intrusion detection mechanisms perform the detection process by identifying the differences between log records. A graph based clustering method was proposed in [21] that clusters similar nodes based on the concept of Euclidean distance and identifies the outliers. Transmission records are used to identify values that correspond to similarity values. A game theoretic based intrusion detection is performed in [22]. Many other methods [23,24] are available, that uses similar distance based techniques for detecting intrusions. These methods are very similar to the classical intrusion detection methods, and do not leverage the complete functionalities of graph databases.Various visualization based methods are also available [26] for effective analysis, but the visualizations are custom built and they cannot be utilized for functionalities other than the defined ones. Graph databases such as Graphviz, Warlus, Neo4j and Gephi are available, using which complex graph operations can be performed on the data, for better data analysis based on parameters which were initially unknown in traditional methods.

3.5. Intrusion Detection using Big Data Analytics

The recent years have seen tremendous researches in big data, which is due to the increase in the information flow in the network. The large amount of traffic (log data) generated from networks (volume) and the speed at which the data is generated (velocity) is sufficient to justify the usage of Big Data technologiesfor the process of intrusion detection. This is a new direction, and research literatures in this area are very less. Hadoop is used as the standard environment for developing Big Data applications. The availability of various algorithms in the Hadoop environment has proved to be a major positive aspect, which attracts researches and researchers towards this technology. The availability of parallelization options in this area is an added advantage. An adaptive detection approach for detecting anomalies using big data is presented in [27], while [28] leverages the parallelization facilities available in the MapReduce to perform effective classification of network data.

Table.3. Summary of IDS/IPS techniques

IDS/IPS technique	Characteristics/ advantage	Limitation/challenges
Signature based detection	• Identifies intrusion by matching captured patterns with preconfigured knowledge base. • High detection accuracy for previously known attacks. • Low computational cost.	• Cannot detect new or variant of known attacks. • High false alarm rate for unknown attacks.
Anomaly detection	• Uses statistical test on collected behavior to identify intrusion. • Can lower the false alarm rate for unknown attacks.	• More time is required to identify attacks. • Detection accuracy is based on amount of collected behavior or features.
ANN based IDS	• Classifies unstructured network packet efficiently. • Multiple hidden layers in ANN increase efficiency of classification.	• Requires more time and more samples training phase. • Has lesser flexibility.
Fuzzy Logic based IDS	• Used for quantitative features. • Provides better flexibility to some uncertain problems.	• Detection accuracy is lower than ANN.
Association rules based IDS	• Used to detect known attack signature or relevant attacks in misuse detection.	• It cannot detect totally unknown attacks. • It requires more number of database scans to generate rules. • Used only for misuse detection.
SVM based IDS	• It can correctly intrusions, if limited sample data are given. • Can handle massive number of features.	• It can classify only discrete features. So, preprocessing of those features is required.

GA based IDS	• It is used to select best features for detection. • Has better efficiency.	• It is complex method. • Used in specific manner rather than general.
Hybrid techniques	• It is an efficient approach to classify rules accurately.	• Computational cost is high.

4. DATA SET ANALYSIS

KDD CUP 99 dataset has been the mostly used dataset in the domain of intrusion detection, due to the fact that it is the only publicly available dataset. It contains both normal traffic and 57 distinct attack vectors. The major drawback of this data set is that it is old and most of the attacks represented here are obsolete and have been fixed. Further, it misses the real background traffic that is necessary for any IDS. Though NSL KDD has been developed with cleaned data from KDD 99, it also has all these downsides.

The next choice for our analysis is the SSENet-2011 dataset [29]. SSENet was constructed using the Tstat tool, and was developed as a collaborative project by five universities with a total of 69 participants. It contains three major classes of attacks; probing attacks, flooding attacks, and privilege escalationattacks along with normal traffic. Table 3 shows the parameters provided in the SSENet dataset, constructed by the Tstat tool.

Table 4: Features Constructed From Tstat Tool

SI. No.	Features
	Network Features
1	Source_IP
2	Source_Port
3	Destination_IP
4	Destination_Port
5	Transport layer protocols(TCP/UDP)
6	Service accessed(HTTP,FTP,SMTP,etc)
7	Number of packets between source and destination
8	Number of segments with ACK bit set
9	No of bytes sent in the payload
10	No. of bytes transmitted in the payload including retransmissions
11	Number of out_of_sequence segments
12	SYN count
13	FIN count
14	Average RTT
15	Standard Deviation RTT
16	Number of retransmitted segments due to timeout expiration
17	Duration of connection in milliseconds
18	Connection type
19	HTTP type(GET/POST)

Connection based features	
20	count_src1: Number of connections made by the same source as the current record in the 100 connections
21	count_dest1: Number of connections made by the same destination as the current record in the 100 connections
22	count_serv_src1: Number of connections with the same service made by the same source as the current record in the last 100 connections
23	count_serv_dest1: Number of connections with the same service made to the same destination as the current record in the last 100 connections

5. PROPOSED RESEARCH FRAMEWORK

The following presents the research framework of our application in intrusion detection and discusses probable technologies that can be used to perform these phases efficiently. Efficiency here is measured in terms of accuracy and speed. It starts with the process of data exploration, which deals with analyzing the data and its contents, that helps in the further phase of data classification. Exploration also provides the user with an idea of the analysis methods that can be used on the data for effective processing. This phase is followed by the data cleaning and then the feature selection. These two phases play a vital role by providing the appropriate data to the user and averts the algorithm from providing biased results. The next three phases, shuffling, segregation and normalization are used for reordering and modifying the data. They are the initial preparation steps for the process of clustering. Shuffling mixes the data such that similar data groups do not end up in isolation in the training or testing datasets. Segregation divides the data for training and testing phases. Normalization, as the name suggests, normalizes the data for processing. Normalization though not compulsory in many methods, if applied can help provide un-biased results. The final phase is the Clustering/ Classification. In case of Classification, the training and test data are used, while clustering uses unlabeled data to provide the results. Table 5 shows effective methods of performing each phase in our intrusion detection system.

Table 5: Proficient Technologies for IDS Phases

Phase	MapReduce (Hadoop)	Graph DB	Machine Learning	Evolutionary Algorithms	Data Mining Algorithms
Exploration	✓	✓			
Cleaning			✓	✓	✓
Feature Selection			✓	✓	✓
Shuffling	✓				✓
Segregation	✓				✓
Normalization	✓				✓
Classification/Clustering	✓	✓		✓	

The proposed framework for IDS will utilize a combination of these methods to provide a hybrid framework for detecting intrusions in the network. The results obtained from analysis are tabulated in Table 4 and the future implementations of IDS frameworks will be based on these results.

6. CONCLUSION

This paper presents a study on the contemporary methods that are used to build Intrusion Detection Systems. Detailed discussions on the components of IDS and IDPS are provided, along with their functionalities. A comprehensive analysis of the methods that perform effective Intrusion Detection is presented. Analysis of these methods, along with tabulations of their specific working modes and accuracy levels are discussed in detail. This paper has been prepared as a part of the author's study in the development of a Hybrid Intrusion Detection System. The results presented here will be analyzed and the technologies that best suits each of the phase willbe identified. These technologies will be analyzed and their appropriate usage areas will be identified for building the proposed Hybrid Intrusion Detection System, which will be our future research direction.

7. REFERENCES

[1] S. Sonawane, Sh. Pardeshi, G. Pradad, A Survey on Intrusion Detection Techniques (Department of Information Technology, Technocrats Institute of Technology, Bhopal, India, April 2012.

[2] N. Jacob, C. Brodley, Offloading IDS Computation to the GPU (Computer Science Department, Tufts University,Medford, 2006.

[3] Vokorokos, Liberios, Michal Ennert, and JánRadušovský. "A survey of parallel intrusion detection on graphical processors." Central European Journal of Computer Science 4.4 (2014): 222-230.

[4] M.S. Clos, A Framework for Network Traffic Analysis Using GPUs (UniversitatPolitecnica de Catalunya, Barcelona, 2010).

[5] Elkan C. Results of the KDD'99 classifier learning contest. SIGKDD. Explor. Newsl 1999;1(2):63e4.

[6] Chen ZF, Qian PD, Chen ZF. Application of PSO-RBF neural network in network intrusion detection. In: Proceedings of the 3rd International Symposium on Intelligent Information Technology Application 2009. p. 362e364.

[7] RG Reynolds. Flocks, herds, and schools: a distributed behavioral model. Computer Graphics 1987; 21(4):25e34.

[8] Ma R, Liu Y, Lin X, Wang Z. Network anomaly detection using RBF neural network with hybrid QPSO. In: Proceedings of the IEEE International Conference on Networking, Sensing and Control 2008b. p. 1284e1287.

[9] Ma R, Liu Y, Lin X. Hybrid QPSO based wavelet neural networks for network anomaly detection. In: Proceedings of the Second Workshop on Digital Media and its Application in Museum and Heritages. 2007. p. 442e447.

[10] Liu H, Jian Y, Liu S. New intelligent intrusion detection methods based on attribute reduction and parameters optimization of SVM. In: Proceedings of the Second International Workshop on Education Technology and Computer Science (ETCS). 2010. P.202e205.

[11] 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 2009. p. 1205e1209.

[12] Wang J, Hong X, Ren R, Li T. A real-time intrusion detection system based on PSO-SVM. In: Proceedings of the International Workshop on Information Security and Application 2009 (IWISA 2009). p. 319e321.

[13] Zhao C, Wang W. An improved PSO-Based rule extraction algorithm for intrusion detection. In: Proceedings of International Conference on the Computational Intelligence and Natural Computing 2009 (CINC '09). p.56e58.

[14] Alipour H, Khosrowshahi E, Esmaeili M, Nourhossein M. ACOFCR: applying ACO-based algorithms to induct FCR. In: Proceedings of the World Congress on Engineering (IWCE) 008. p. 12e17.

[15] Abadeh MS, Habibi J. A hybridization of evolutionary fuzzy systems and ant colony optimization for intrusion detection. The ISC International Journal of Information Security 2010; 2(1):33e46.

[16] Ramos V, Abraham A, ANTIDS: Self organized ant based clustering model for intrusion detection system. In: Proceedings of The Fourth IEEE International Workshop on Soft Computing as Trans disciplinary Science and Technology (WSTST'05) 2005. p. 977e986.

[17] Tsang W, Kwong S. Unsupervised anomaly intrusion detection using ant colony clustering model. In: Proceedings of the 4th IEEE International Workshop on Soft Computing as Trans disciplinary Science and Technology 2005. p. 223e232.

[18] Tsang CH, Kwong S. Multi-agent intrusion detection system in industrial network using ant colony clustering approach and unsupervised feature extraction. In: Proceedings of the IEEE International Conference on Industrial Technology 2005 (ICIT 2005). p.51e56.

[19] Feng Y, Zhong J, Ye CY, Wu ZF. Clustering based on self organizing ant colony networks with application to intrusion detection. In: Proceedings of the Sixth International Conference on Intelligent Systems Design and Applications (ISDA '06). 2006. P.1077e1080.

[20] Y. Zhang, W. Lee, Intrusion detection in wireless ad-hoc networks, in: The 6th Annual International Conference on Mobile Computing and Networking, Boston, MA, USA, 2000, pp. 275–283.

[21] D. P. Jeyepalan, E. Kirubakaran ,(April 2013),"A Novel Graph Based Clustering Approach for Network Intrusion Detection", International Journal of Computational Intelligence and Information Security, Vol. 4 No. 4,ISSN: 1837-7823.

[22] D. P. JeyepalanE. Kirubakaran,(2014), "A Co-operative Game Theoretic Approach to Improve the Intrusion Detection System in a Network using Ant Colony Clustering", International Journal of Computer Applications,Volume 87 - Number 14.

[23] Marinakis, Yannis, et al., (2011), "A hybrid ACO-GRASP algorithm for clustering analysis." Annals of Operations Research 188.1: 343-358.

[24] Ganapathy, Sannasi, et al., (2013), "Intelligent feature selection and classification techniques for intrusion detection in networks: a survey." EURASIP Journal on Wireless Communications and Networking 2013.1: 1-16.

[25] Aziz, Amira Sayed A., and Aboul Ella Hassanien. "Multilayer Machine Learning-Based Intrusion Detection System." Bio-inspiring Cyber Security and Cloud Services: Trends and Innovations. Springer Berlin Heidelberg, 2014. 225-247.

[26] Luo, Bin, and Jingbo Xia. "A novel intrusion detection system based on feature generation with visualization strategy." Expert Systems with Applications 41.9 (2014): 4139-4147.

[27] Zhang, Ji, et al. "Detecting anomalies from big network traffic data using an adaptive detection approach." Information Sciences (2014).

[28] Chen, Tieming, et al. "Efficient classification using parallel and scalable compressed model and its application on intrusion detection." Expert Systems with Applications 41.13 (2014): 5972-5983.

[29] Vasudevan, ARi, E. Harshini, and S. Selvakumar. "SSENet-2011: a network intrusion detection system dataset and its comparison with KDD CUP 99 dataset." Internet (AH-ICI), 2011 Second Asian Himalayas International Conference on. IEEE, 2011.

[30] P. Fanfara, A. Pekár, Usage of Hybrid Honeypots an Intrusion Detection System Mechanism, SCYR 2012: Proceedings from conference : 12th Scientific Conference of Young Researchers, 2012

[31] V. Marinova-Boncheva, A Short Survey of Intrusion Detection Systems (Institute of Information Technologies, Sofia, 2007)

[32] Kolias, Constantinos, Georgios Kambourakis, and M. Maragoudakis. "Swarm intelligence in intrusion detection: A survey." *computers & security* 30.8 (2011): 625-642.