Real Valued Negative Selection for Anomaly Detection in Wireless Ad Hoc Networks

Azri Abdul Majid [a], Mohd Aizaini Maarof [b]

Group on Artificial Immune System and Security (GAINS)
Department of System and Computer Communication
Faculty of Computer Science and Information System
Universiti Teknologi Malaysia
81310 Skudai, Johor
Malaysia
Tel : +607-5532398   Fax : +607-5565044

E-mail : [a] azri5@siswa.utm.my, [b] maarofma@fsksm.utm.my

Abstract

Wireless ad hoc network is one of the network technologies that have gained lots of attention from computer scientists for the future telecommunication applications. However it has inherits the major vulnerabilities from its ancestor (i.e., the fixed wired networks) but cannot inherit all the conventional intrusion detection capabilities due to its features and characteristics. Wireless ad hoc network has the potential to become the de facto standard for future wireless networking because of its open medium and dynamic features. Non-infrastructure network such as wireless ad hoc networks are expected to become an important part of 4G architecture in the future. In this paper, we study the use of an Artificial Immune System (AIS) as anomaly detector in a wireless ad hoc network. The main goal of our research is to build a system that can learn and detect new and unknown attacks. To achieve our goal, we studied how the real-valued negative selection algorithm can be applied in wireless ad hoc network network and finally we proposed the enhancements to real-valued negative selection algorithm for anomaly detection in wireless ad hoc network.

1. Introduction

Wireless ad hoc network has gained its popularity among computer security researchers because of its open medium and dynamic features. However it does not just introducing new vulneribilities, but also inherits the major vulnerabilities from its ancestor, i.e., the fixed wired networks. But it cannot inherit all the conventional intrusion detection capabilities due to its features and characteristics. Wireless ad hoc network has the potential to become the *de facto* standard for future wireless networking. Infrastructureless network such as wireless ad hoc networks are expected to become an important part of 4G architecture in the future [1].

There are two type of intrusion detection; host-based and network-based. As mentioned above, wireless ad hoc network is a purely non-infrastructured network. Obviously, fixed wired network intrusion detection system (i.e., misuse or anomaly detection) cannot operate well without major modification in this new wireless environment. New or enhanced intrusion detection system (IDS), especially the anomaly detection model, design, approach, technique and architecture need to be studied for this purpose.

Due to the wireless ad hoc dynamic features, anomaly detection has been known as the best technique for intrusion detection. However, anomaly detection IDS has to undergo long training sessions of noiseless data to construct fairly accurate statistics for nomal condition. Another issue of anomaly detection is it relies on data belonging to one single class or limited instances of some known class with the goal of detecting all unknown class [2]. Particularly, a major difficulty in using machine learning methods for anomaly detection lies in the making the learner discover boundaries between known and unknown classes.

Eventhough anomaly detection is the only intrusion detection technique that is best suited with wireless ad hoc networks, the false alarms is still become the major issue. Particularly, the data collection is very important. It becaused the training data is very crucial and sensitive. In

infrastructure networks, anomaly detection and data collection can be done at server, switch, gateway or router. But in wireless ad hoc network, each node must act both as terminals and information relays and participates actively in the routing protocol. There is no centre point for data collection in wireless ad hoc networks.

The human immune system is a highly dynamic, distributed, adaptive and decentralized system which can recognizes and classifies many different and unseen pathogens (i.e., non-self cells) such as bacterias and viruses at any given time [3]. Negative selection is the process for self and non-self discrimination occurred in thymus gland in our body where antibodies can be produced using only the self cells information. In artificial immune system, negative selection algorithm is one of the components which commonly used in the field of fault tolerant and fault detection, intrusion detection and several others applications.

This paper is organized as follows: In section 2, we explain the features and characteristics of wireless ad hoc networks and how the wireless ad hoc network differs from fixed wired network. Then in section 3 and 4, negative selection and real-valued negative selection algorithm is explained and discussed. An enhancement for real-valued negative selection algorithm was proposed as anomaly detector in wireless ad hoc network. Section 5, is the conclusion to the discussion of this selected issues.

2. Wireless Ad Hoc Networks

Wireless ad hoc network is a collection of wireless nodes that can be rapidly deployed as a multi-hop packet radio network without the aid of any existing network infrastructure [4]. Each node in wireless ad hoc network is a switch and router at the same time. Wireless ad hoc network also recognized with its fully decentralized topology. Communication beyond the transmission range is possible by having all nodes act both as terminals and information relays [5]. This is known as hopping which can increase ad hoc network scalability [6]. Individual nodes will discover dynamically which other nodes they can communicate with. Ad hoc networks is very ideal for military, save-and-rescue operation or in a situation where installing infrastructure network is not possible or when the purpose of the network is too transient or even for the reason that the previous infrastructure network was destroyed [4, 7].
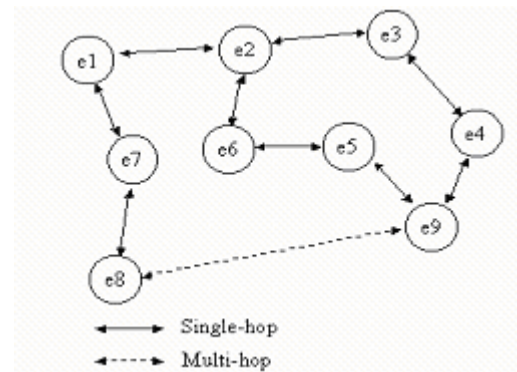
Figure 1. Example of Small Scale Wireless Ad Hoc Network

Figure 1 shows the example of small scale wireless ad hoc network consists of nine nodes. All the nodes were connected to form a network. Single-hop is a term used when a node can directly connect to other node to communicate. And multi-hop is implemented when several nodes has to become both a router and a relay to other node to form a connection. One of the distinct characteristics of wireless ad hoc network is that all participating nodes have to involve in the routing process. Nevertheless, routing in wireless ad hoc network needs to rely on the trustworthiness of all nodes that are participating in the routing process. Thus, traditional routing protocols designed for fixed wired network cannot be applied in ad hoc network [4].

Its dynamic features and decentralized topology makes security in wireless ad hoc network is very hard to achieve. Existing intrusion detection system for infrastructure network is not suitable for wireless ad hoc network. Obviously, wireless ad hoc network have its own vulnerabilities that cannot be solved by existing wired security solutions [4]. It becomes even worst when attacks on wireless ad hoc networks can come from any direction and it can be targeted to any node [7]. The most important things in ad hoc networks are its routing process where all nodes need to involve and participate in a very cooperative way. Once the node has been hijacked, the entire wireless network can be paralyzed by disseminating false routing information [7].

Wireless ad hoc networks have a very bright future in several fields of application. Especially in the field of military operation [8], search-and-rescue mission, academic environment or anywhere where file sharing is very important and at the any location and situation where network infrastructure cannot be found. It also very cost effective when compared to the fixed wired network due to its features and characteristics.

2.1 Artificial Immune System

The main function of the Immune System is to protect the body against different types of pathogens, such as harmful viruses, bacteria and parasites [9]. It consists of a large number of different immune cells which interact in order to provide the detection, elimination of the pathogens and finally recovered (heal) the body from the attacks.

Artificial Immune System is an approach where natural human immune system is translated into component of computer algorithms. The most widely used of AIS algorithms are Negative Selection Algorithm, Clonal Selection Algorithm, and Immune Network Algorithm. These algorithms have evolved into versions of enhanced algorithms. However, only negative selection algorithm will be discussed in this paper. The advantages and disadvantages of this algorithm was identified, and then modified to make it works with anomaly detection in wireless ad hoc networks.

2.1.1 AIS and Wireless Ad Hoc Network

The important and most difficult step in applying the AIS concept into our application is the mapping from AIS to our field of domain [10]. Wireless ad hoc network architecture operates in open medium through limited range of radio signal.
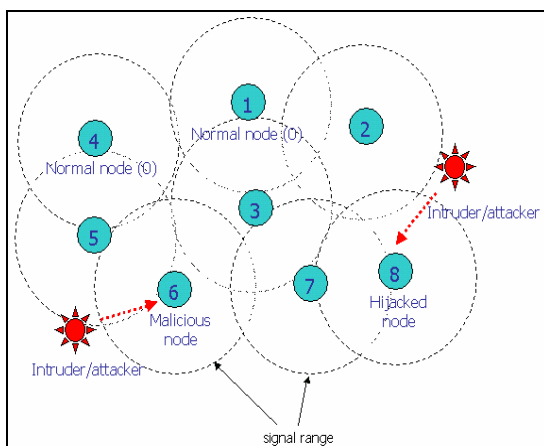

Figure 2. The Architecture of Wireless Ad hoc Network

Figure 2 illustrates the architecture of wireless ad hoc network which consists of eight nodes. Each node can communicate directly in peer-to-peer connection. The radio signal used in wireless ad hoc network can be considered as the chemical or signal released by immune cells such as B-cell or T-cell to interact with each other in natural human immune system. Implementing this onto wireless ad hoc networks using Artificial Immune System

approach is very challenging task. Three factor need to be considered when doing the mapping process:

a. data collection
b. normal profiling
c. classification

The type of data to collect as the input data is very important because it will effects the normal profiling stage. And in the final stage, classification will rely on the normal profiled condition to compare between the observed condition and test condition for intrusion detection.

3. Negative Selection Algorithm

Negative selection algorithm was first introduced by Forrest [11]. The algorithm has two phases which is 1) generate a set of detectors and 2) monitor the protected data (i.e, matching function). There are two major issues raised by this algorithm. The issues are meaningless information and scalability. Meaningless information occured when negative selection has to perform its processing (i.e matching function) in binary value form. In the end, negative selection has to retrieve the result (output) to its original form as in input. This is known as low level representation of the detectors and also the main cause of low meaningful of domain knowledge [12]. However, this binary valued negative selection also was applied firstly in wireless ad hoc by [5] as misbehavior detection. Eventhough negative selection has shown a promising result in misbehavior detection, its false alarm is still high.

		Binary			Gray		
	r	ND	D%	FA%	ND	D%	FA%
r-contiguous	7	0			40	3.96%	1.26%
	8	343	15.84%	16.84%	361	16.83%	16.67%
	9	4531	53.46%	48.48%	4510	66.33%	48.23%
	10	16287	90.09%	77.52%	16430	90.09%	75.0%
	11	32598	95.04%	89.64%	32609	98.01%	90.4%
r-chunk	4	0			2	0.0%	0.75%
	5	4	0.0%	0.75%	8	0.0%	0.75%
	6	18	3.96%	4.04%	22	3.96%	2.52%
	7	98	14.85%	16.16%	118	18.81%	13.13%
	8	549	54.45%	48.98%	594	73.26%	47.47%
	9	1942	85.14%	72.97%	1959	88.11%	67.42%
	10	4807	98.01%	86.86%	4807	98.01%	86.86%
	11	9948	100%	92.92%	9948	100%	92.92%
	12	18348	100%	94.44%	18348	100%	94.44%
Hamming	12	1	0.99%	3.03%	7	10.89%	8.08%
	13	2173	99%	91.16%	3650	99.0%	91.66%
	14	29068	100%	95.2%	31166	100%	95.2%
Rogers & Tanimoto	9/16	1	0.99%	3.03%	7	10.89%	8.08%
	10/16	2173	99%	91.16%	3650	99%	91.66%
	11/16	29068	100%	95.2%	31166	100%	95.2%
	12/16	29068	100%	95.2%	31166	100%	95.2%

Table 1. The Result of Detection using different Matching Function in the Negative Selection Algorithm [13].

Table 1 shows the result of detection using different macthing function techniques using Mackey-Glass data series. Rogers and Tanimoto represent the euclidean distance technique which produces the good result in true detection. However, the false alarm is still high.

In negative selection algorithm, false alarm is due to the scalability issue when the negative selection has to generate sufficient detectors to ensure that its detection capability can detect large numbers of anomaly traffic. The issue here is how much detectors has to be generated and is sufficient enough in order to ensure its good detection capability. If the amount of generated detectors is too small, detection will be not effective and will produce high volume of unwanted false alarm. Otherwise, if the detectors are too large, it could be unmanageable and can caused bad effects for wireless ad hoc node itself. This is known as energy consumption issue.

Negative selection algorithm which uses binary representation for its matching rules has raised several issues such as scaling problems, meaningless data representation and sharp distinction between self and non-self boundary [14]. To overcome these problem, an enhanced negative selection algorithm was introduced by [15,16,17,18,19,20,21]. It was known as real-valued negative selection algorithm.

4. Real-Valued Negative Selection

We believe that anomaly detection in wireless ad hoc networks should use real-valued negative selection because it will use only normal samples to generate abnormal samples.

Real-valued negative selection algorithm was first proposed by Fabio Gonzalez in his publications [15,16]. The idea of using real-valued is to use representation schemes that are closer to the problem space. This higher-level representation provides advantages such as increased expressiveness, i.e, the possibility of extracting high-level knowledge especially from the generated detectors.

This algorithm operates on a unitary hypercube $[0,1]^n$. A detector $d = (c, r_{ns})$ has a center $c \in [0,1]^n$ and a non-self recognition radius $r_{ns} \in ¡$. An every self element $s = (c, r_s)$ has a center of self radius r_s . Self radius was used to consider other elements as self which are close to the self center. If an element lies within a detector, then it will classified as non-self, otherwise as self. An element e lies within a detector $d = (c, r_{ns})$ if the euclidean distance $dist(c,e) = \left(\sum_{i=1}^{n} (c_i - e_i)^2 \right)^{1/2} < r_{ns}$.

Euclidean function was primarily used as the matching function in the real-valued representation. Candidate detectors are generated randomly. As in the negative selection algorithm, those that match any self samples will be eliminated.

In the mean time, real-valued negative selection is based on heuristic rules that try to move the detectors away from each other and from the self point. An indirect result of this is an increase in the non-self space covered by the set of detectors.
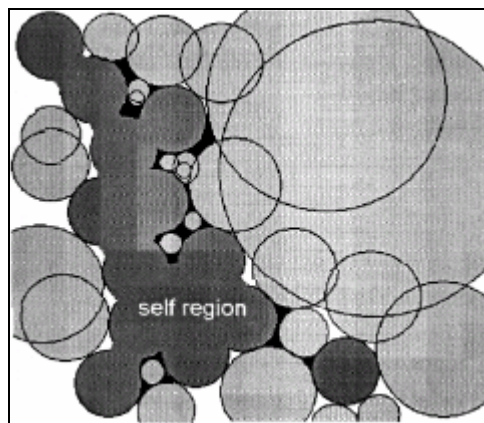
Several version of real-valued negative selection has evolved such as randomized real-valued negative selection RRNS using the mathematical models which uses mathematical modelling such as Monte Carlo simulation technique as its randomized detector generation engine [17].

Real-valued negative selection with variable sized detectors known as V-detector was proposed by [19]. It was a new scheme of detector generation and matching mechanism for negative selection algorithm. The V-detector algorithm can evaluate and estimate the coverage automatically when the detector set is generated. It randomly determines the center of a detector which must not lie within the boundary of the hypersphere of a self element. Then the radius is dynamically resized until the boundary of a region comes in contact with a self element.
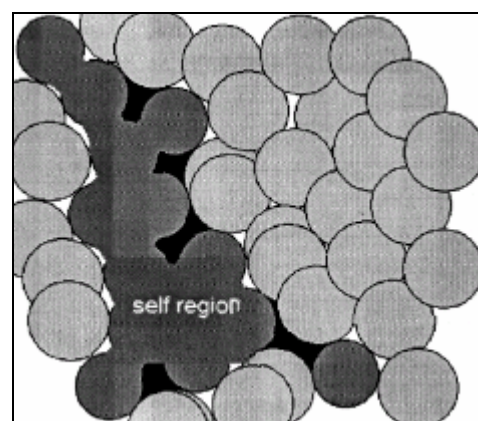
(a) Constant-sized detectors

(b) Variable-sized detectors

Figure 3. Main Concept of Negative Selection and V-detector [19]

Algorithm	Detection Rate		False Alarm Rate		# Detectors or # Support Vectors	
	Mean	SD	Mean	SD	Mean	SD
V-detector$^{r_s=0.000005}$	2.66	8.35	0.00	0.00	1.37	0.52
V-detector$^{r_s=0.00001}$	2.40	7.12	0.00	0.00	1.36	0.51
V-detector$^{r_s=0.00005}$	1.75	6.05	0.00	0.00	1.39	0.56
V-detector$^{r_s=0.0001}$	1.58	5.73	0.00	0.00	1.33	0.50
V-detector$^{r_s=0.05}$	1.21	4.59	0.00	0.00	1.48	0.59
V-detector$^{r_s=0.1}$	0.65	3.46	0.00	0.00	1.59	0.67
Self-Detector$^{r_s=0.000005}$	100	0.00	0.00	0.00	9727	0
Self-Detector$^{r_s=0.00001}$	100	0.00	0.00	0.00	9727	0
Self-Detector$^{r_s=0.00005}$	100	0.00	0.00	0.00	9727	0
Self-Detector$^{r_s=0.0001}$	100	0.00	0.00	0.00	9727	0
Self-Detector$^{r_s=0.05}$	100	0.00	0.00	0.00	9727	0
Self-Detector$^{r_s=0.1}$	99.99	0.02	0.00	0.00	9727	0
ocSVM$^{\nu=0.005}$	99.78	0.03	0.05	0.02	55.70	1.56
ocSVM$^{\nu=0.01}$	99.82	0.02	0.99	0.02	103.40	1.50
ocSVM$^{\nu=0.05}$	99.87	0.02	4.95	0.03	491.15	1.27
Parzen-Window$^{u=0.005}$	99.93	0.02	0.00	0.00	—	—
Parzen-Window$^{u=0.01}$	99.93	0.02	0.00	0.00	—	—
Parzen-Window$^{u=0.05}$	99.93	0.02	0.00	0.00	—	—

Table 2. Classification Results for Self Detector Classification using KDD Dataset by [21]

Figure 3 (a) illustrates the main concept of negative selection using the constant-sized detector while figure 3 (b) is the negative selection using the variable-sized detectors. The area in dark gray indicates the self region and the light gray is the detector coverage. And the black color indicates the holes which detector coverage is impossible. This means, there is certain criteria that detection cannot be done. But using variable-sized detectors can decrease the holes area.

Stibor [21] has concluded that negative selection whether it is a binary scheme or real-valued has raised the algorithm complexity issue. He has proposed the Real-valued Self Detector Classification which based on positive selection approach. The main goal of the algorithm is to overcome the scaling problem inherent in the hamming shape space negative selection algorithm. There is no non-self detectors exist. This means that no detector generation phase is needed. But the classifcation for each unseen element is computationally expensive [21].

Table 2 shows the result of classification result of intrusion detection using several version of real-valued negative selection algorithm. Self detector shows the most promising detection result. But as mentioned above, it was computionally expensive.

Eventhough real-valued negative selection has used the representation schemes that are closer to the problem space, it has to consider on the three important parameters namely self radius, estimated coverage and maximum numbers of detectors. Before we implement this algorithm in the wireless ad hoc networks, we must answer the following research question:

- What size of radius is the best for self and non-self to cover the self and non-self coverage?

- What size of estimated coverage need to be covered by the detectors to get a high detection rate?

- What maximum numbers of detectors need to be generated until we can have a good detection capability?

These question is important solve the anomaly detection problems in anomaly detection for wireless ad hoc networks using artificial immune system approach.

4.1 Anomaly Detection using Real-Valued NSA

Intrusion Detection technique can be divided into misuse detection and anomaly detection. Misuse detection will detect signature or pattern of known attacks. The great advantage of misuse detection is it can produce high rate of true detection. However, it cannot detect new attack pattern and it have to update its attack pattern very often. In self-organized such as wireless ad hoc networks, this is not impossible but very difficult to implement. Anomaly detection is the technique that best suited with

wireless ad hoc dynamic environment. It can detect new attack pattern without rely upon previous attack signature or patterns.   The major drawback is it produces high percentage of false alarm. Besides, it also has to go for a long training session to produce its normal condition before attacks can be detected.

Several researches have applied anomaly detection technique in wireless ad hoc environment such as in [7, 9, 10,22].   And the major drawback of their anomaly detection is the high rate of false alarm. We believed that the main reason is because the selection criteria of the input features is not appropriate enough to conclude what is   actually the normal condition in wireless ad hoc networks.

Obviously, [9, 10, 23, 24] has shows that the real-valued representation scheme must be used instead of binary representation in applying anomaly detection in wireless ad hoc environment. Thus, real-valued negative selection is the most appropriate approach for this reason.

5. Conclusion

Our study has shows that Artificial Immune System (i.e. Real-valued Negative Selection Algorithm) is still relevant to be applied in wireless as hoc networks as an anomaly detector. But, several modifications need to be done to the original algorithm to make it significant with this new dynamic environment. Especially to its affinity measurement (i.e. matching function), knowledge representation and mapping from AIS to wireless ad hoc networks.

Anomaly detection in wireless ad hoc networks should be simple but effective enough to detect intrusions. Training session should be shorter and it must not computationally expensive since energy usage in wireless ad hoc networks is very limited.

We also believe that there is no sharp distinction between normal and anomaly condition. In this case, we are considering on adapting fuzzy logic approach in our real-valued negative selection algorithm to tackle the false true detection and false alarm issue.

References

[1] Chlamtac I., Conti M., Liu J., (2003) "Mobile Ad hoc Networking: Imperatives and Challenges" Ad-Hoc Networks Journal, Inaugural Issue, No.1. Vol. 1, 2003.

[2] Fan, W., et.al., (2001), "Using Artificial Anomalies to Detect Unknown and Known Network Intrusions." Proceedings of the 2001 IEEE International Conference on Data Mining, 29 November - 2 December 2001, San Jose, California, USA. Pp 123-130.

[3] Stibor T., Mohr P., (2005),"Is negative selection appropriate for anomaly detection" Genetic And Proceedings of the 2005 conference on Genetic and evolutionary computation (GECCO 2005), Washington DC, USA, Pages: 321 – 328, ISBN:1-59593-010-8

[4] Stamouli I., (2003), "Real-time Intrusion Detection for Ad hoc Networks.", The University of Memphis : PhD Dessertation.

[5] Le Boudec J. Y., Sarafijanovic, S. (2004), "An Artificial Immune System Approach to Misbehavior Detection in Mobile Ad-Hoc Networks", Proceedings of Bio-ADIT 2004 (The First International Workshop on Biologically Inspired Approaches to Advanced Information Technology), Lausanne, Switzerland, January 2004, pp. 96-111.

[6] Unnikrishnan. P., (2004), "Introduction And Analysis of DSR Protocol." Routing - from baseline to state-of-the-art Seminar on Internetworking, Helsinki University of Technology, Spring 2004

[7] Zhang, Y et.al. (2000), "Intrusion Detection in Wireless Ad Hoc Networks." Proceedings on Mobile Computing and Networking (MobiCom '2000) August 6-11, Boston, Massachussets.

[8] Stamouli I., Argyroudis P., and Tewari H., (2005), "Real-time Intrusion Detection for Ad hoc Networks", Proceedings of the 6th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), Taormina, Italy, 14-16 June, 2005, pp. 374-380

[9] Boudec L. J. ; Sarafijanovic S. (2003), "An Artificial Immune System Approach to Misbehavior Detection in Mobile Ad Hoc Networks." Technical report No. IC/2003/59, September 2003.

[10] Le Boudec J. Y., Sarafijanovic, S. (2003), "An Artificial Immune System Approach with Secondary Response for Misbehavior Detection in Mobile Ad-Hoc Networks.", Technical report No. IC/2003/65, November 2003.

[11] Forrest S., et. al, (1994) "Self-Nonself Discrimination in a Computer." In Proceeding if 1994 IEEE Symposium on Research in Security and Privacy, Los Alamitos, CA: IEEE Computer Society Press, pp. 202-212.

[12] Gomez, J., Gonzalez, F., Dasgupta, D., (2003), "Combining Negative Selection and Classification Techniques for Anomaly Detection." Fuzzy Systems, 2003. FUZZ '03. The 12th IEEE International Conference on , Volume: 2 , 25-28 May 2003 Pages:1219 - 1224 vol.2.

[13] Gómez J., González, F., and Dagupta D., (2003), "An Immuno-Fuzzy Approach to Anomaly Detection." In Proceedings of The IEEE International Conference on Fuzzy Systems, St. Louis, MO, May 2003.

[14] González F., Dagupta D., and Gómez J. (2003) "The Effect of Binary Matching Rules in Negative Selection", In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO)*, July 11-16, 2003

[15] Gonzales, F. (2003), "A Study of Artificial Immune System Applied to Anomaly Detection." The University of Memphis: PhD Dessertation.

[16] Gómez J., González F., and Dasgupta D., (2003), "An Immuno-Fuzzy Approach to Anomaly Detection." In Proceedings of The IEEE International Conference on Fuzzy Systems, St. Louis, MO, May 2003.

[17] González F., D. Dasgupta and J. Gómez, (2003), "A Randomized Real-Valued Negative Selection Algorithm", In *Proceedings of the 2nd International Conference on Artificial Immune Systems*, Edinburgh, UK, September 2003.

[18] González F., Dasgupta D., (2003), "Anomaly Detection Using Real-Valued Negative Selection" Journals *Genetic Programming and Evolvable Machines*, 4(4), pages 383-403, Kluwer Acad. Publ., December 2003.

[19] Ji Z., Dasgupta D., (2004) "Real-valued Negative Selection Algorithm with Variable-Sized Detectors" Proceedings of Genetic and Evolutionary Computation Conference (GECCO-2004) Seattle, Washington June 26-30, 2004.

[20] Dasgupta D., Kumar K., et. al., (2004) "Negative Selection Algorithm for Aircraft Fault Detection" 3rd International Conference on Artificial Immune Systems Catania, Sicily.(Italy) September 13-16 2004.

[21] Stibor T., Timmis J., and Eckert C. (2005), "A Comparative Study of Real-Valued Negative Selection to Statistical Anomaly Detection Techniques" International Conference on Artificial Immune Systems (ICARIS), Banf, Canada. LNCS 3627. pp. 262-275 C.Jacob Et. al. (Eds) 2005.

[22] Zhang. Y et. al (2003), "Intrusion Detection Techniques for Mobile Wireless Networks." ACM Klumer Wireless Networks Journal Vol 9, No. 5, pp. 545-556, September 2003.

[23] Nadkarni, K. and Mishra, A (2004) "A Novel Intrusions Detection Approach for Wireless Ad Hoc Network." Wireless Communications and Networking Conference, 2004. WCNC. 2004 IEEE , Volume: 2 , 21-25 March 2004
Pp:831 – 836.

[24] Mishra, A. Nadkarni, K. Patcha, A (2004), "Intrusion Detection in Wireless Ad Hoc Networks" Wireless Communications, IEEE ,Feb 2004, pp: 48-60Volume: 11,   Issue: 1 ISSN: 1536-1284.