A Cooperative Negative Selection Algorithm for Anomaly Detection

Praneet Saurabh
Deptt.of CSE, TIT
Bhopal, M.P, India

Bhupendra Verma
Deptt.of CSE, TIT (E)
Bhopal, M.P, India

ABSTRACT
Artificial Immune System (AIS) is a convoluted and complex arrangement derived from biological immune system (BIS). It possesses the abilities of self-adapting, self-learning and self-configuration. It has the basic function to distinguish self and non-self. Negative Selection Algorithm (NSA) over the years has shown to be competent for anomaly detection problems. In the past decade internet has popularized and proliferated into our lives immensely. Internet attack cases are increasing with different and new attack methods. This paper presents a Cooperative Negative Selection Algorithm (CNSA) for Anomaly Detection by integrating a novel detector selection strategy and voting between them to effectively identify anomaly. New introduced mechanisms in CNSA enable it to cover more self region correctly and efficiently. It also reduces computational complexities. Experimental results show high anomaly detection rate with less false positive alarm and low overhead in most of the cases.

Keywords
Artificial Immune System, Biological Immune System, Negative Selection Algorithm, Anomaly

1. INTRODUCTION
Anomaly detection is a primary and essential part of modern life. It's concern and importance remains same regardless of the domain. A general approach for anomaly detection is to define a range of deviation for each parameter, and if the deviation of parameter lies outside the range, it is abnormal otherwise normal [1, 3]. Pervasiveness of computers and importance of information lures hackers to take advantage of the vulnerabilities. Anomaly content and its methodologies have evolved and become more complex and very difficult to recognize the anomalous behavior.

Artificial Immune System (AIS) is inspired from Biological Immune System (BIS) [7, 10]. AIS models the processes and procedures of BIS which helps an organism survive under various challenging circumstances [14]. AIS is recognized as a new evolving paradigm for solution of complex computational problems. AIS is used to find solution in the domain of anomaly detection, fault diagnosis, computer security, and problems of optimization [2, 9]. Reaction and response of BIS against new and complex threats motivates researchers and scientists to model anomaly detection after BIS.

AIS models various procedures of BIS, out of those Negative Selection Algorithm (NSA) and Clonal Selection Algorithm are the prominent ones [14]. Negative selection is based on the procedure that goes inside the thymus to select T-cells. It is based on immune system's skill to differentiate between unknown antigens (non-self) and normal (self); without reacting to the normal (self) [11]. In other way it performs anomaly detection. NSA is widely used in anomaly detection domain and quite efficient with the results. All these features make NSA an exciting prospect and caught attention of research fraternity [12]. The initial Clonal selection concept was proposed by Burnet in 1959 and is based on generating those cells that are capable to identify antigens. T cells in biological immune system when encounters an antigen stimulate B cells which have the power to match antigens. Once these B cells are stimulated they clones themselves to produce new ones. These new clones are not exactly same but they undergo mutation to identify various new pathogens. Afterwards it becomes memory cells and geared up to respond to same pathogen if encountered in future [14]. Negative Selection Algorithm and Clonal Selection Algorithm offer many features that can be exploited to address anomaly detection problem.

Detectors in NSA aim to cover more and more self and non self space correctly. Detector generation and selection is very important as these detectors forms base for anomaly detection [5, 14]. Many previous works focused on it's optimization but still it needs to be addressed properly. This paper presents Cooperative Negative Selection Algorithm (CNSA) for Anomaly Detection which integrates novel detector selection strategy into the training phase and voting in the next phase between detectors to identify an anomaly correctly. Experimental results show that these mechanisms helps Cooperative Negative Selection Algorithm (CNSA) identify new and unseen anomalies. Section 2 contains related work in anomaly detection and AIS in anomaly detection. Section 3 presents Cooperative Negative Selection Algorithm (CNSA) for Anomaly Detection. Result and analysis with discussion is covered in Section 4, Section 5 concludes this paper.

2. RELATED WORK
2.1. Anomaly Detection System
Anomaly can be defined as which is not normal, it represents the behavior that can be termed as abnormal. A general approach for anomaly detection defines a range of deviation for each metric, and if the deviation of metric lies outside the range, it is abnormal otherwise normal [1, 3]. Due to so much of importance different anomaly detection schemes were proposed which includes statistical [3], machine learning and data mining. Unfortunately, these proposed techniques faces difficulty in many scenarios of anomaly detection for computer security [9]. Potential reasons are these are based on collecting, analyzing and extracting evidences after an attack which often makes it slow and lack of availability of abnormal samples at training phase. Furthermore lack of self-learning and self-adapting qualities reduces its efficiency.

2.2. Biological Immunity motivated approaches
In the recent years various biological immunity motivated approaches have been proposed to successfully carry out anomaly detection [4, 5, 11]. The task of anomaly detection is analogous to Biological Immune System (BIS) as both have the

task to distinguish normal and abnormal which violates the defined policies [3].

Biological Immune System (BIS) protects all organisms against different attacks and threats. It is has the great potential of pattern recognition which identifies foreign cells (non-self) and prevents it from entering body [7, 8]. Immune system is acquainted with various characteristics such as distinctive, autonomous, acknowledgment of aliens, distributed, and noise tolerant. Central lymphoid generates immune cells which are matured in thymus. Matured and useful T-cells are released while discarding all the remaining ones. These matured T-cells recognize antigens and subsequently takes suitable and appropriate measures [6]. The immune system is broadly divided into two parts: innate and adaptive immune system. The innate system is the defense shield with which an individual is born made of complement system and phagocytes. The adaptive immune system is mainly composed of white blood cells also termed as acquired immunity because it builds a memory over a period of time to achieve a faster response when the same threat or antigen is confronted next time [8].

2.3. Artificial Immune System

Artificial Immune System (AIS) is encouraged from Biological Immune System (BIS) used for solving complex computational problems [7]. Apart from attractive qualities of AIS such as self-configuration, self-learning, self-adaptation and distributed coordinating, its ability to identify self and non self fascinates research fraternity the most. Lot of research work in AIS is centered around three theories: negative selection, clonal selection and immune theory [14]. New theories are developed and introduced to address different problems of various domains.

2.4. Negative Selection Algorithm & Clonal Selection

Negative Selection Algorithm (NSA) revolves around the most interesting feature of self/non-self identification, pioneered by Forrest and her group [13]. NSA revolves around the theory of T-cell maturation in thymus and self tolerance of immune system. NSA generates detectors and eliminates those ones that detect self. This results in a group of detectors that has the potential to detect non-self. Different variations in NSA have been suggested over a period of time from its inception to solve problems of various domains including anomaly detection, fault detection and optimization [5, 13, 14].

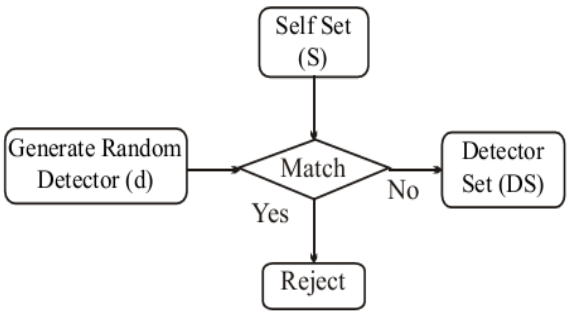

Fig.1 Detector Generation & Selection

Clonal Selection theory illustrates creation of immune cells when activated in the presence of an antigen. T cells when encounters an antigen stimulate B cells which are capable of matching antigens. The moment these B cells are stimulated, they initiate the process to produce clones of themselves. These clones are not the precise copy of the previous ones. Actually they undergo mutation to enhance their uniqueness to address and match different antigens [10]. This B cell become specific and becomes memory cell to facilitate adaptive immune response so that it can react faster to the same antigen if encountered in future [14].

2.5. Representation of Negative Selection Algorithm

Artificial immune system (AIS) entices researchers with its attractive qualities such as self-configuration, self-learning, self-adaptation, and distributed coordinating. The most prominent functionality is its ability to distinguish self and non-self. A lot of research and landmarks has been achieved since Jerne introduced it. New AIS models have been proposed time and again to solve different kinds of problems from the domain of computer security, data mining, clustering, data analysis, and classification. (NSA) is represented broadly through two representations: binary representation (low) and real valued representation (high). Many work in NSA used binary representation to present the problem because it provides finite problem space which makes the analysis of problem easier and straightforward. But it is marred with certain limitations that include lack of scalability and limited information extraction. Gonzalez et. al. introduced real valued negative selection (RNS) in quest to prevail over the limitations of binary representation. Real valued representation includes advantages of better eloquence and more high level knowledge extraction from generated detectors [5].

2.6. NSA in Anomaly Detection

In the beginning NSA was developed to detect change in the system [9]. Negative selection approach is frequently used in the anomaly detection domain because of astonishing similarities between requirements of problem space and features offered by NSA. BIS very successfully identifies foreign invaders and this feature is exploited to develop anomaly detection system. Forrest in 1994 [13] introduced first theory with binary strings. Later on anomaly detection system using real valued representation with variable-sized detectors (V-detector) was developed [5]. V-detector uses variable-sized detectors for anomaly detection. Detectors in NSA cover self and non-self space to recognize anomaly in the problem space [5, 11]. Variability in the self samples and random detector generation and selection creates the problem of holes or uncovered space represented in figure 2. Also a detector can identity an anomaly by just one comparison with test data which many a times increases false positive. Furthermore detection generation and selection is search and optimization problem and still it needs to be realized properly.

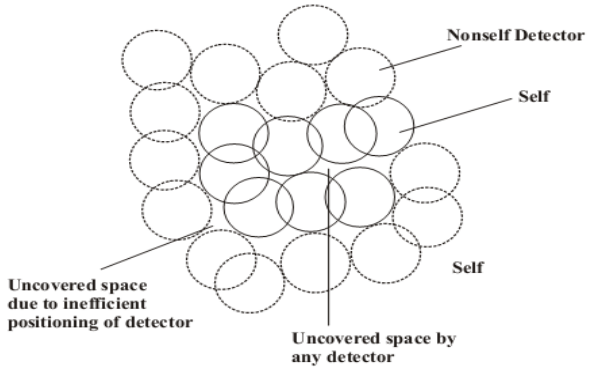

Fig.2 Problem of holes or uncovered space

3. COOPERATIVE NEGATIVE SELECTION ALGORITHM (CNSA)

Detectors in proposed Cooperative Negative Selection Algorithm (CNSA) aim to maximize coverage of the non self region and minimize false coverage of self. At the detector selection stage CNSA employs an aggressive strategy to select efficient detectors.

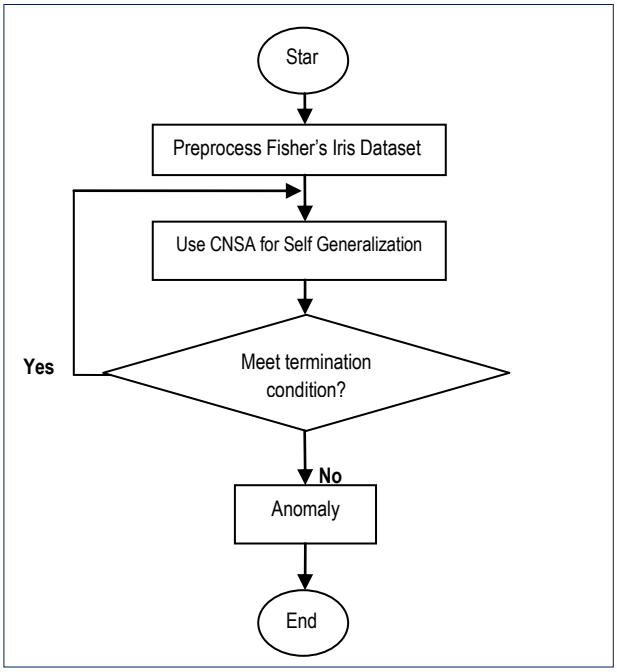

Fig.3 Cooperative Negative Selection Algorithm (CNSA)

3.1. Detector Generation and Selection

Earlier work in NSA generated detectors of small or of the same size to cover self/ non-self space due to this some space remained uncovered. A few overlapping cases are also recognized. These existing methods gave rise to problem of similarity and reflected in low detection rate of anomaly and high false alarm rate. In this proposed approach random real valued number generator generates random unique numbers. Detectors are generated on the basis of these numbers. Non similar value of detector (*d*) in shape space represents and establishes its uniqueness. In the proposed CNSA detectors are fine tuned so that it is not discarded after just one comparison from the training set (*TS*). Finetuning of detectors enables it to efficiently and correctly cover more self non-self space. Detector coverage is an optimization problem and many works have focused on it.At the detector selection step detector (*d*) is compared with the instances of *Test Set* (*TS*) and Euclidian similarity (*D*) is calculated illustrated in (*I*).

$$D(dx_i, TSy_i = \sqrt{\sum_i (dx_i, TSy_i)^2} = \quad \text{----- (I)}$$

$$D(dx_i, TSy_i) < B\_T \quad \text{---- (II)}$$

If the Euclidian similarity (*D*) between detector (*d)* and Training Set (*TS*) is less than Binding Threshold (*B_T*) then in normal scenario detector (d) is discarded, but in CNSA detector (*d*) isnot discarded but it is fine tuned by a factor T_V. Detector (*d*)is altered by a factor of *'V'*, till equation (*II*) becomes affirmative. This reflects detector (d) has now become dissimilar and selected. Finetunig helps CNSA in taking less iterations for detector selection and henceforth it reduces with computational complexities.

Apart from this uniqueness a new parameter detector power (d_P) is introduced in CNSA. Power of detector (d_P) is a relative parameter and its final value is determined when all the detectors are compared with all the instances of training set (*TS*). At the time of generation of detector (*d*) power of detector 'd_P' remains 0. In the next step detector (d_P) is matched with the instances of training set (*TS*). Detector Power (d_P) is increased by one after each negative match of detector (*d*) with training set (*TS*). Detectors having power greater than (P_T) power threshold are selected for Detector Set (*DS*). This mechanism helps in selection of better and efficient detectors for anomaly detection.

3.2. Cooperation between detectors for Anomaly Detection

Different detectors cooperate and coordinate between themselves to flag any instance or pattern in the test set (*TeS*) as an anomaly.

Algorithm for anomaly detection with CNSA

```
Input: RS = Rule Set
Input: DS= Detector Set
Input: TeS= Test Set,
Input: A_T= Affinity Threshold=0
Input: V_T= Vote Threshold=0
ASSIGN DS and RS
for all  (DS, RS) 1 to N for all packets in TeS 1 to k
CALCULATE D= EUCLIDIAN DISTANCE (x_i, d)
if D >A_T
DS_K detects TeS as Anomaly
    increase Vote by 1
  end if
  if (Vote >=V_T)
    ANOMALY
  else
    NORMAL
  end if
end for
```

Detectors vote among themselves for a particular test set instance. Each successful comparison of detector and test set instance increases the vote count by 1. If the total vote for that instance is greater than vote threshold (V_T) then only that instance is termed and flagged as an anomaly otherwise normal.

A new variable Risk Count (*R_C*) is introduced to monitor longevity of detectors (*d*) and detector set (*DS*). If the risk count of detector set (*DS*) increases beyond *R_C* then these *DS* are discarded. Clonal selection clones to create new detectors. These detectors forms new detector set (*DS*) New cloned Detector Set (*DS*) are equipped with new detectors (*d1, d2,..,d*$_n$)containing more relevant and efficient detectors to detect more and new attacks correctly and efficiently.

3.3. Anomaly Detection by CNSA

Cooperative Negative Selection Algorithm (CNSA) integrates attribute of detector power (d_P) at detector selection stage to select powerful detectors. These powerful detectors forms detector set (*DS*). Detectors ($d_1, d_2, \ldots$) from Detector Sets ($DS_1, DS_2 \ldots$). are compared with the samples of Test Set (*TeS*). Each Detector Set (*DS*) uses its own detectors ($d_1, d_2, d_3, \ldots\ldots d_n$) to detect the anomalies of the incoming Test Set (*TeS*). Euclidean similarity measure (*D*) is calculated between detector (*d*) of

Detector Set (*DS*) and *1* to k^{th} element of *TeS*, test sample illustrated in equation (*II*).
Equation (*II*) represents that if the similarity measure (*D*) is greater than Affinity_Threshold (*A_T*) then this test sample is marked as an anomaly temporary. As a result Detector (*d*) increases vote count '*V*' for this sample of test set (*TeS*).

$$D(dx_i, TeSy_i) = \sqrt{\sum_i (dx_i, TeSy_i)^2} = \|x_i, y_i\| \quad \text{----- (III)}$$

$$f_{match}(d_i, TeS, A\_T) = \begin{Bmatrix} Vote++; if \left||d_i - x|\right| \le A\_T \\ Vote--; if \left||d_i - x|\right| > A\_T \end{Bmatrix} \text{--(IV)}$$

Information about vote count of test sample with a vote count *'V'* is shared with other detectors. Other detectors in Detector Set '*DS*' also compares the test sample and contributes to vote count *'V'*. This process is repeated till all the *Detector Sets*(*DS*) test the test sample presented by equation ***(IV)***. If the final vote-*'V'* for a packet increases over the *Vote* Threshold (V_T), then the packet is marked as an anomaly. The new features introduced in *CNSA* helps in selecting efficient and powerful detectors. Detector Set having powerful detectors detects anomalies more proficiently. Cooperation between detectors of detector set reduces false positive.

4. EXPERIMENTAL STUDY

Cooperative Negative Selection Algorithm (CNSA) introduces finetuning, power of detectors (d_P) and voting mechanism between detector set (*DS*) to cover self and non-self space more efficiently. Detection Rate and False Alarm is evaluated to determine the proficiency of anomaly detection system. Detection rate is the number of anomalies detected out of the available ones, while False alarm rate represent how many normal is identified as an anomaly. Any system with high detection rate and low false alarm is considered to be good. Detection Rate (*DR*) is calculated by: $DR = \frac{TP}{TP+FN} * 100$;where TP means True Positive and FN stands for False Negative. Higher detection rate is always desired. False Positive Rate is obtained by $FPR: \frac{FP}{TN+FP} * 100$, where FP means false positive and TN represents true negative. Lesser false positive means that the system is properly detecting self as self.

4.1. Dataset

All the experiments of both CNSA and RNS are carried on Fisher Iris Dataset [16] and then comparisons have been drawn. It consists of 50 samples from each of three species of Iris flowers (Setosa, Virginica and Versicolor). Four features were measured from each sample; they are the length and the width of sepal and petal. RNS with constant sized detectors cannot adapt to the diversity of self and nonself space. Also it cannot build an appropriate profile of the system. All experiments are performed on three classes of Fisher Iris Dataset. Test Set (*TS*) is composed 40 instances out of 50 in each case of out of the three classes Setosa, Virginica and Versicolor. Testing Set (*TeS*) contains 10 remaining instances of the same class and all the instances of the remaining two classes. This arrangement of training and test set verify self and non-self coverage. Detector power (*dP*) is 0.7; means the detectors who failed to match 70 out of 100 test samples were discarded. All the results are the average of 50 runs on the same configuration and then Detection Rate and False Positive Rate is calculated. A good detection system labors to achieve high detection rate and low false positive rate.

4.2. Detector Selection vs Rejection

All the experiments are performed on Fisher's Iris dataset and then comparison between CSNA and RNS is drawn under varying Affinity Threshold. Training Set (*TS*) is composed of one variety of Iris flower. Test Set (*TeS*) contains some instances of the same variety of Iris flower and the two other species. Same variety of flower sample here represents self and other variety represents non-self. Test Set (*TeS*) has both normal and anomalous instances. Detectors of both CNSA and RNS are trained by 100% of dataset. This experiment is carried with the intension to identify the difference which fine tuning makes in detector selection. Table 1 illustrates the effect of finetuning in selection of detectors. These results are an average of 50 runs of different number of detector generation while tuning factor (T_V) remains as 1. Detectors are generated with variation in the training sample size of preprocessed data by using both RNS and the proposed CNSA.The results in table 1 very clearly reflect that the selection percentage of detectors by using the proposed CNSA is much higher than RNS. As the training data increases from 25% to 100 % then difference in selection percentage of RNS and CNSA increases very sharply because larger amount of self helps in creating self profile efficiently.

Table. 1

Training Data	Method	Detector Selection (%)	Detector Rejection (%)
25%	RNS	76.38	23.62
	CNSA	96.15	3.85
50%	RNS	42.6	57.4
	CNSA	96.47	3.53
100%	RNS	19.64	80.36
	CNSA	99.9	0.1

Finetuning introduced in detector selection aids this high selection percentage as detectors are nor discarded after just one comparison but it is tuned so that it become non similar. As the training samples increases the selection percentage of detectors in RNS comes down and detector rejection percentage increases since RNS discards a detector just after one failed comparison. As a result it has to generate more detectors to map the whole self sample.

4.3. Detection Rate and False Positive

Table 2 and Figure 4 illustrate the performance of RNS and CNSA when affinity threshold is varied between 0.5 to 0.6. Results are an average of 50 runs. CNSA attain better detection rate for different data sample (Setosa, Verginica and Versiculor). All the detectors are trained by 100% of self samples. Finetuned and Powerful detectors in detector set cover self and nonself space properly and correctly which makes CNSA detect more anomalies and achieve high detection rate. Cooperation between detectors helps CNSA lower down False Positive as a test sample (*TeS*) can only be labeled as anomaly when it's vote count becomes greater than vote threshold. (V_T). False positive of CNSA remains lesser in all the cases which highlight the significance of cooperation between detectors introduced in CNSA.

Table: 2

Training Data	Affinity Threshold	Method	Detection Rate	False Positive
Setosa	0.5	RNS	95.1	2.7
	0.5	CNSA	98.3	1.4
	0.6	RNS	97.2	3.6
	0.6	CNSA	99.7	1.5
Virginica	0.5	RNS	96.6	2.4
	0.5	CNSA	98.2	1.2
	0.6	RNS	97.3	3.3
	0.6	CNSA	99.7	1.4
Versicolor	0.5	RNS	95.4	2.1
	0.5	CNSA	98.7	1.1
	0.6	RNS	97.5	3.4
	0.6	CNSA	99.8	1.2

These results also show that affinity threshold as an important parameter for detecting anomaly. Figure 4 demonstrates the curves of RNS and CNSA under different affinity thresholds. It is quite easy to figure out that CNSA achieves better detection rate and lower false positive than RNS.

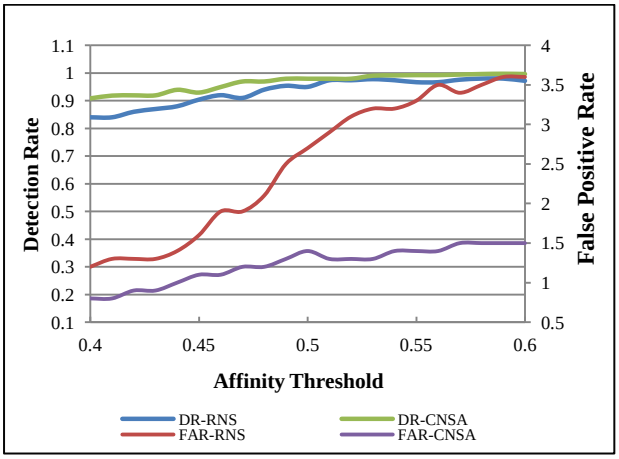

Fig. 4 Affinity Threshold vs Detection Rate vs FPR (Setosa)

All the results in Table 2 and Fig.4 points out that CNSA accomplishes better performance as compared to RNS without powerful and cooperative detectors.

4.4. Effect of training data size

Observation in table: 3 demonstrate how CNSA fares when trained with different samples of training set (*TS*). Detectors of CSNA are trained with different sizes of test set.

Table.3

Training Data	Data Size	Method	Detection Rate	False Positive
Setosa	50	CNSA	98.1	1.1
	100	CNSA	99.7	1.5
Virginica	50	CNSA	98.5	1
	100	CNSA	99.7	1.4
Versicolor	50	CNSA	98.6	1.3
	100	CNSA	99.8	1.2

Detection rate and False positive rate ate calculated using same test set (*TeS*).Affinity threshold (*A_T*) and detector power (d_P) remains 0.6 and 0.7. Against variation in training sample the relative average detection rate and false alarm of CNSA is calculated.Limited and less size of training data means limited information is available for building self profile. Information about the profile is very important in case of recognizing self and nonself. Results point out that CNSA very successfully builds self profile. CNSA when trained with 50% of the training data (*TS*), it achieves a detection rate of 99.1 % and false positive rate of 1.1. Detection rate increases to 99.4 and false positive rate comes down to 0.52 when CNSA is trained on 100% of training set. Other two results of the two samples Verginica and Versiculor are inline with Setosa. CNSA successfully builds self profile even if training data is limited and achieves high detection and false positive rate while maintaining stability.

5. CONCLUSION

Cooperative Negative Selection Algorithm (CNSA) aided by finetuned and powerful detectors cooperates among themselves to attain better results. Finetuned and Powerful detectors facilitates and enhances detection rate as detector set only contain those detectors which is more dissimilar and has high chance of identifying an anomaly. CNSA also successfully builds self profile irrespective of the available knowledge at training stage. Cooperation between detectors plays very important role in lowering false positive rate as any pattern is not labeled as an anomaly till it gets substantial votes from other detectors. Moreover CNSA remains stable under all the scenarios. Experimental result establishes the fact that the introduced mechanism makes CNSA much more comprehensive and improved as it CNSA identifies self and non self quite effectively. This powers CNSA to identify anomalies more proficiently and correctly.

6. REFERENCES

[1] B. Mukerjee, L. T. Heberlein, K. N. Levitt, "Network Intrusion Detection", IEEE Network, Vol. 8, No.3, 1994, pp 26-41.

[2] Charles Cresson Wood, "The Human Immune System as an Information Systems Security Reference Model", Computers and Security, Elsevier Vol.6, 1987, pp- 511-516.

[3] D. E. Denning, "An Intrusion-Detection Model", IEEE Transactions on Software Engineering, Vol. 13, No. 2, February 1987, pp. 222-232.

[4] F. Esponda, S. Forrest, P.Helman, "A Formal Framework for Positive and Negative Detection schemes", IEEE Transactions on System, Man and Cybernetics, Vol. 34, No. 1, 2003, pp 357-373.

[5] F.A.Gonzalez, D.Dasgupta, "Anomaly Detection Using Real-Valued Negative Selection", Genetic Programming and Evolvable Machine,Vol.4. No.4, 2003, pp. 383-403.

[6] Hiroyuki Nishiyama and Fumio Mizoguchi, "Design of Security System Based on Immune System", pp 138-143, IEEE 2001.

[7] Mark Burgess, "Biology, Immunology and Information security", Information Security Technical Report, Science Direct, Vol.12, 2007, pp 192–199.

[8] N. K. Jerne, "Towards a network theory of the immune system", Ann. Immunol.(Paris), 125C, 1974, pp. 373-389.

[9] P.Saurabh, B.Verma, S.Sharma, "Biologically Inspired Computer Security System: The Way Ahead", SNDS, Communications in Computer and Information Science, Vol, 335, Springer, 2012, pp. 474-484.

[10] Richard E. Overil, "Computational immunology and anomaly detection", Information Security Technical Report, Science Direct, Vol.12, 2007, pp 188-191.

[11]S.Forrest, A. S. Perelson, L. Allen, R. Cherukuri, "Self-nonself discrimination in a computer", IEEE Symposium on Research in Security and Privacy, 1994, pp. 202–212.

[12]S.Forrest, S.A. Hofmeyr, A.Somayaji, "Computer Immunology," Communications of the ACM, Vol. 40, No.10, 1997, pp. 88–96.

[13]S.Forrest, S.A.Hofmeyr, A.Somayaji, T.A.Longstaff , "A sense of self for Unix processes", Proceedings IEEE Symposium on Security and Privacy, 1996, pp 120-128.

[14]S.Ramakrishnan, S.Srinivasan,."Intelligent agent based artificial immune system for computer security—a review", Artificial Intelligence Review, Vol. 32, No. 1-4, December 2009, pp 13–43

[15] W.Wang, X.Guan, X.L.Zhang, Processing of massive audit data streams for real-time anomaly intrusion detection", Computer Communications, Vol. 31, No.1, 2008, pp.58–72

[16] StatLib-datasets archive. http://lib.stat.cmu.edu/dataset/