Model and Algorithm in Artificial Immune System for Spam Detection

Ismaila Idris

Dept of Cyber Security Science, Federal University of Technology, Minna, Nigeria.

idris.ismaila95@gmail.com

Abstract

A spam detection model based on negative selection algorithm is proposed in this paper. The artificial immune system creates techniques to solve complex computations, aiming to developing immune based models. This is done by distinguishing self from non-self. Preliminary mathematical analysis will expose the computation and experimental description of the method and how it is applied to spam detection. A new detector model and matching rule model are also generated for effective matching of both self and non-self in other to burst the detector performance of the model. Our unique matching technique use in the negative selection algorithm help the model to overcome the limitation of a normal negative selection algorithm in defining harmfulness of self and non-self. This improves the requirement of the model and satisfactory requirement in terms of true positive and false positive rates. The experimental result confirms that the proposed model is able to establish a better true positive on an unknown spam

.

Keywords

Artificial immune system, Negative selection, Computer security, Algorithm, Model

1.Introduction

Models and application in artificial immune system is coming up as an active and attractive field of great diversity. Great source of inspiration to computational model are drawn from the already knowledge of the immune system. Over the past years, rapid expansion of computer network system as change the world. It is essential for an effective computer security system because attacks and criminal intend are increasingly popular in computer network[1] . There are several measure put in place by many companies in the area of creating anti spam software based on signatures and have a very efficient performance in detecting spam fast. Though, new variation of spam and unknown spam are very difficult to detect by this software. The traditional way of detecting spam based on signature is no more efficient for today systems. Recent years, researchers are interested in the field of immune system in achieving computer security. Immune theory was first applied to computer abnormality detection by Forrest et al. After then, several proposition in the area of immune system as made tremendous success.

Negative selection algorithm, while not reacting to the self cells uses the immune system capability to detect unknown antigens. Its mechanism protects body against self reactive lymphocytes. Receptors are made through a pseudo-random genetic re-arrangement process during the generation of T-cells [2]; they then undergo a censoring process in the thymus called the negative selection. In this process T-cells that do not bind to self-proteins are destroyed. Therefore, immunological function and protection of the body against foreign antigens is possible through circulation of matured T-cells. Artificial negative selection algorithm was proposed to follow that pattern; generating random detectors and then discard those that match self samples. It is a spam detection in which training data from only one of the classes are available. With these illustrations, the traditional negative selection algorithm believes that all self are not dangerous and all non- self

DOI : 10.5121/ijaia.2012.3107

are dangerous. In our approach to this, the computer security system needs to only recognize dangerous spam instead of classifying all non-self to be dangerous. Unlike the traditional method, the extraction of spam and non-spam from the training set in [3] are been added together, which help in the reduction of false positive. More so, the negative selection algorithm with penalty factor in [4] penalizes the non-self by the use of penalty factor instead of deleting the non-self. It rather kept them in a library. In our proposed approach. Spam and non-spam detectors shall be utilized successively in other to achieve best false positive rate. Our model shall eliminate waste of non-self that cannot match with self by creating a uniform platform for both self and non-self to be able to operate has a detector instead of discarding non-self that those not match self. This will eliminate false negative and improve the detector performance.

2. Related Work

Several artificial immune algorithm are been created with imitation of Clonal selection theory. The Clonal selection principle comprises of antigen recognition and differentiation in the memory cell. Selection, cloned and mutation takes place through the selection of some of the good antibody during iteration session. This resulted into new antibody, were the best among them merges with the original population while worst antibody of the previous generation are been substituted with randomly generated new ones. The principle of this theory is to ensure that the immune system is able to keep interconnected B cells network for the recognition of antigen. The stability of the network was due to the connectivity of these cells in several ways. If the affinity of two B cells exceed a threshold, it is said to be interconnected, were there strength is directly proportional to its shared affinity. Early works of [5] and [6] has provided many avenue for researcher to explore as many model are been developed in its literature. The ability of the immune system been utilized to detect antigen that are not known to respond to self cells is the Negative selection mechanism; while protecting the body over self reacting lymphocytes. Receptors are created through pseudo-random genetic rearrangement procedure during the generation of T-cell, which then go through a censoring procedure in the thymus; this process is known as negative selection. With this procedure, response against self protein by T-cells is destroyed while only those that were able to match to self protein are given the chance to leave the thymus. This T-cell that are allowed is then lunched in to the body protecting it against antigens.[7]. Negative selection algorithm was proposed by [7]. The principle behind it is to create a set of detectors that could be use to detect malware. This was achieved by randomly creating candidates and those that recognize training self data are been discarded. In furtherance of our related work, we are going to give a summary of some of the existing work of the different AIS techniques and models in this section with focus on the last five years.

[8] proposed an adaptive Clonal algorithm for optimal phasor measurement unit (PMU) placement. In this proposition, hyper mutation probability and recombination operators of the circle supplement population of the CLONALG algorithm are adjusted. With this, the optimization process is improved and optimal traps are avoided. A better version of OPT-IA was also introduced by [9] called opt-IMMALG. The introduction of new inversely proportional hyper mutation operator and the changing of binary string with a real-coded one are the main improvement of the algorithm. An improved Clonal selection algorithm that is focus on CLONALG with a unique mutation methods was presented by [10], called the self-adaptive chaotic mutation. The new algorithm uses the logistic chaotic sequence to create the first antibody population which is the major improvement of the algorithm, though the hyper mutation uses self –adaptive chaotic mutation. [11] also presented a parallel Clonal selection algorithm that is use in problem solving of Graph Coloring , uses an island model where all processor are able to work on their antibody pool to enhance its ability. An Immune Based Network Intrusion detection System (AINIDS) was proposed by [12]. This consists of five components: a data collector, a packet head parser and feature extraction, antibody generation and antigen detection, co-stimulation with

report, and rule optimization components. The creation of an algorithm called aiNet to solve function optimization problems was also proposed by [13]. Also a Local Network Neighborhood Artificial Immune System (LNNAIS) model for data clustering was also proposed by [14]. LNNAIS uses the concept of artificial lymphocyte (ALC) neighborhood to ascertain the linking network between the artificial lymphocytes. The network lack affinity threshold in its model that will ascertain if two artificial lymphocytes could be link together to create a network. The individual indexes ascertain the lymphocyte neighbor and they also make interaction and learn from each other in other to have a good representative of patterns. Tree Structured Artificial Immune Network (TSAIN) for data clustering and classification was also proposed by [15].There is no need setting a threshold for these model as topological link was set up among two antibodies immediately when one of them has produced for the other. Other proposed immune network algorithms are also introduced in [16].

Another earlier work that proposed artificial immune system is [17], whose work mainly base on immune network model; though, did not use negative selection algorithm but its model of learning has a flavor of negative selection from positive example. The matching rule differs from the rcb in some ways, though it uses binary representation. In [17], it was pointed out that the threshold of matching is selected to be half of the antibody, this procedure may not be applicable to a wide range of application; though it is alright for every application and also very reliable biologically. A review and comparison studies of five negative selection algorithm was done by [18]; Linear, Greedy, Exhaustive, Binary template and NSmutation. The time and space complexity that are alright for comparison of the five algorithms are examine. [7] originally proposed the exhaustive detection generation [18] which present the basic concept of negative selection. An hybrid model for phrase chunking with artificial immune system and rule base technique was proposed in [21]. Also, a negative selection algorithm with penalty factor in [4] uses penalty factor to penalize the non-self instead of deleting it. It was rather kept them in a library. In this paper, we shall be looking at Spam and non-spam detectors been utilized effectively in other to achieve best false positive rate. Our model shall create a memory space for both self and non-self to enable the both be used as a detector instead of eliminating non-self that those not match. This will help in drastically reduce the false positive and false negative for efficient detector.

3. Negative Selection Based Algorithm

The negative selection algorithm definition defines self to be equal to the collection of element in a feature space U.

U is represented by list of features which corresponds to the space of states of a system where
S=Subset of space that are considered as normal for the system.
R= Set of detectors generated.

(1)

$R \neq S \rightarrow R$ fails to match any string in S (2)

This approach analyzes happening in Negative immune system by generating random detectors and discard those that match any element in the self set. Continually matching S for changes with detectors R against S if S ever matches R. Change is known to have occurred as detectors are not suppose to match any string in S.

Negative selection while not reacting to the self cells uses the immune system capability to detect unknown antigens. Its mechanism protects body against self-reactive lymphocytes. Receptors are made through a pseudo-random genetic rearrangement process during the generation of T cells, and then they undergo a censoring process in the thymus called the negative selection. In this process T-cells that do not bind to self-proteins are allowed to leave the thymus while those that react against self-proteins are destroyed [19]. Therefore, immunological functions and protection of the body against foreign antigens is possible through circulation of matured T-cells. The algorithm is as stated below.

4. Problem Definition

The higher the possibility of a spam, when the probability is higher in calculating the probability of occurrence of a given value. In other methods, models are built to predict the future behavior of systems or processes base on recent and the formal states. In this scenario, a malware alarm is raised if the normal states of the system differ from the predicted state. Generally, malware is considered as a deviation from a set of a normal states with assumption of distance in this space that allows for measure of deviation [20].

Below is the negative selection of malware detection model problem to be address in this work. In identifying the state of a system as a self or non self is the essence of spam detection problem definition. This is represented by a set of features as shown below:

Let's make vector of the feature as the set of the system which is the system state space.

$a^i = (a_1^i\text{---------------}a_b^i) \in [0,1]^b$ (3)

Each state of the feature been represented by a set

$\cup \subseteq [0,1]^b$ (4)

This includes the feature vectors which correspond to all the probability state of the system.

$\cup$ = Universal set of all the system. (5)

0 and 1 represent system being a self or non-self. (6)

For normal subspace (crisp characterization) set of feature vectors self$\subseteq\cup$, which indicate a non-spam. Therefore non-self which is its compliment defined as non-self = u – self, where non-self indicate spam in the system. Using its characteristics function, we can define self or non-self as follows

$a_{self}: [0,1]^b \longrightarrow [0,1]$ (7)

$a_{self}: \overrightarrow{(a)} \begin{cases} 1 \; if \; \vec{a} \in self \\ 0 \; if \; \vec{a} \in non-self \end{cases}$

(8)

The above is characterized by artificial immune system were cells are distinguish from foreign antigens. The term self represent cells in the immune system and non-self represent the foreign antigens in the system. Though, there are difference between the spam and the non spam state.

Where as normal subspace (non-crisp characterization) features of a spam and non spam is extended to pick values of intervals [0,1]

$\beta^{self} : [0,1]^a \rightarrow [0,1]$ (9)

This value represent degree of either it is a spam or not a spam. Where by 1 indicates that it is not a spam and 0 indicates that it is a spam. The intermediate value represent element with some degree of being a spam or non spam. Also, binary decision as to be in cooperated. It is simple to go from non-crisp characterization to the crisp one by creating a limitation.

$\beta self$, L $(\vec{a}) = \begin{cases} 1 \; if \; \beta \; self \; (\vec{a}) > L \\ 0 \; if \; \beta self(\vec{a}) \leq L \end{cases}$ (10)

With the sample $self^I \subseteq$ self, a good estimate of the normal space can be created with characteristic function. a^{self} in the crisp characterization case and $\beta \; self$ In the non crisp characterization case. This function should be able to tell if there is spam or not

5. Proposed Spam Detection Model

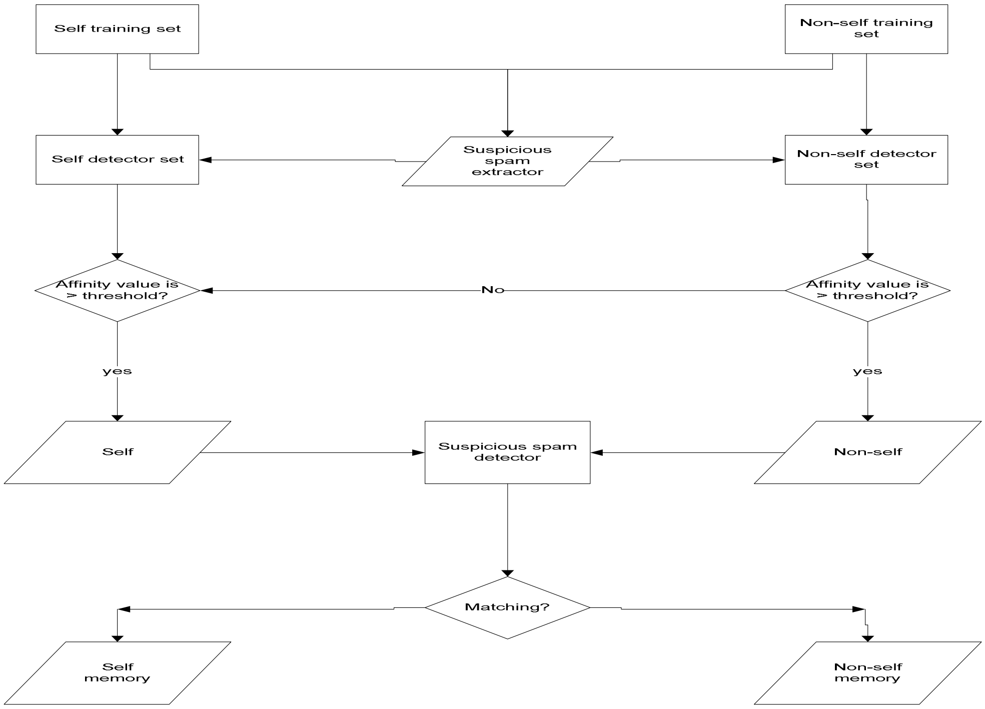

Fig 1. Proposed detector model

The proposed model as shown above in figure1 consists of suspicious spam extractor and a suspicious spam detector as shown in the architecture of the design model. In the suspicious spam extractor; both the self detector set and the non-self detector set are extracted from the spam and non-spam programs of the training data set after creating the suspicious spam detector. After suspicious program (spam and non-spam) in the suspicious spam detector using the suspicious spam extractor, an r matching rules that was initiated will be computed between the suspicious program and the memory self and memory non-self. The proposed matching algorithm focused essentially on memory self and memory non-self, the classification of the gene in antibody variable region enable the combination of the memory-self and memory-non self in different region. If the matching value exceeds the given program classification threshold, we classify the program as spam; otherwise it will be considered as non-spam

5.1 Matching Rule Model

The proposed matching rule is as stated below:
After the division of the data set in to training and testing data set; our training data set will be further divided in to self detector and nonself detector (x,y) respectively using the suspicious spam extractor.

Affinity (x,y) $= \sum match$ (x,y)/l, / x /= l (11)

x = Affinity between memory-self x

y = Affinity between memory-non self y

l = length of the memory-self x and the memory nonself y

$$\text{Match }(x,y) = \begin{cases} 1, & \sum_{i=0}^{n} match\left(x \rightarrow x_{i,} y\right)/l,/\ x\ /-\ l \\ 0, & otherwise \end{cases} \tag{12}$$

$$\text{Match }(x,y) = \begin{cases} 1, & \sum_{i=0}^{n} match\left(y \rightarrow y_{i,} x\right)/l,/\ y\ /=\ l \\ 0, & otherwise \end{cases} \tag{13}$$

Where, the Matching of x is the total match value between memory-self x

And the Matching of y is the total match value between memory- non-self y

The dataset is given as:

S= {<x, y, affinity>|x$, y \in \bigcup_{i-1}^{n} H,$ affinity ∈N} (14)

Affinity is the ability of the self to match with both self and non-self

This exists when memory-self and memory-non self are generated.

The set is divided in to self training set and non-self training set

Let TR denotes training set

$$\text{TR }(x) = \begin{cases} \emptyset; \quad t = 0 \\ TR(t-1) + TRnew\ (t), t \geq 1 \end{cases} \tag{15}$$

TRnew (t) = {x∈AG, f_{match}(y.(x)} (16)

TRnew (t) = {y∈AG, f_{match}(x.(y)} (17)

Therefore; for threshold value of x and y we have

$$f_{match}(y) = \begin{cases} 1, f_{affinity}(y)\ /\ l > \alpha \\ 0, \quad otherwise \end{cases} \tag{18}$$

$$f_{match}(x) = \begin{cases} 1, f_{affinity}(x)\ /\ l \geq \alpha \\ 0, \quad otherwise \end{cases} \tag{19}$$

α Represent the threshold value.

Also, ability of self matching self and non-self we have:

$$f_{affinity}(x) = \max\,(y_{1,}y_{2}\ \ldots\ldots\ldots\ldots, y/\ l - l_x\ /+1) \qquad (20)$$

$$f_{affinity}(y) = \max\,(x_{1,}x_{2}\ \ldots\ldots\ldots\ldots, x/\ l - l_y\ /+1) \qquad (21)$$

$$x_i = \sum_{i=1}^{\min\,(l,l_y)} \theta_{ij} \qquad (22)$$

$$y_i = \sum_{i=1}^{\min\,(l,l_x)} \theta_{ij} \qquad (23)$$

$$\theta_{ij} = \begin{cases} 1, & x_i = y_{i+j\_1}, 1 \le i \le / l - l_y / +1\ , 1 \le j \le l \\ 0, & otherwise \end{cases} \qquad (24)$$

$$\theta_{ij} = \begin{cases} 1, & y_i = x_{i+j\_1}, 1 \le i \le / l - l_x / +1\ , 1 \le j \le l \\ 0, & otherwise \end{cases} \qquad (25)$$

Where,

$f_{matc\square}$ is the match function of self and non-self

$f_{affinity}$ is the affinity function of self and non-self

L is the length of self and non-self

If the affinity between self and non-self is greater than the threshold value $\propto$, then $f_{matc\square}$ returns to 1, otherwise it returns to 0; more so, the greatest number between self and non-self is returned by the function $f_{affinity}$ vice versa.

Match is the set that consist of self who match both self and non-self
TRnew is the new memory-self and memory-non-self set that is composed of the antibody whose affinity with self is larger than the threshold value β or affinity with non-self larger than the threshold value $\propto$. Therefore, to detect spam, a matching algorithm is presented below.

Basically, an memory-self and memory-non-self is use to detect spam. If a self or non-self matches a memory-self or memory-non-self affinity greater than the threshold value, we recognize it has spam.

$$f_{affinity}(x) = \max\,(y_{1,}y_{2\ldots\ldots\ldots\ldots}, y\ /\ l - l_x\ /\ +1) \qquad (26)$$

$$f_{affinity}(y) = \max\,(x_{1,}x_{2\ldots\ldots\ldots\ldots}, x\ /\ l - l_y\ /\ +1) \qquad (27)$$

Where,

$$x_i = \sum_{i=1}^{\min(l,l_y)} \theta_{ij} \tag{28}$$

$$y_i = \sum_{i=1}^{\min(l,l_x)} \theta_{ij} \tag{29}$$

$$\theta_{ij} = \begin{cases} 1, & a_{i=} y_{i+j-1,} \quad 1 \le i \le l_a - l_y/+1, 1 \le j \le l_a \\ 0 & otherwise \end{cases} \tag{30}$$

$$\theta_{ij} = \begin{cases} 1, & b_{i=} x_{i+j-1,} \quad 1 \le i \le l_b - l_x/+1, 1 \le j \le l_b \\ 0 & otherwise \end{cases} \tag{31}$$

Where, a and b are both memory-self and memory-non-self whose length are l_a and l_b respectively.

l_y and l_x are the lengths of the detected files.

The bigger the affinity, the higher the match value between memory-self, memory-non-self, self and non-self, and the more possibility of self or non-self were a spam.

The function match is finally given as

$$f_{matc\square} \text{ (b,x)} = \begin{cases} 1, & f_{affinity} (b,x) / l_b \ \ge \propto \\ 0, & otherwise \end{cases} \tag{32}$$

$$f_{matc\square} \text{ (a,y)} = \begin{cases} 1, & f_{affinity} (a,y) / l_a \ \ge \propto \\ 0, & otherwise \end{cases} \tag{33}$$

Where $\propto$ is the threshold value ranging from 0 to 1

5.2. Proposed Matching Architecture

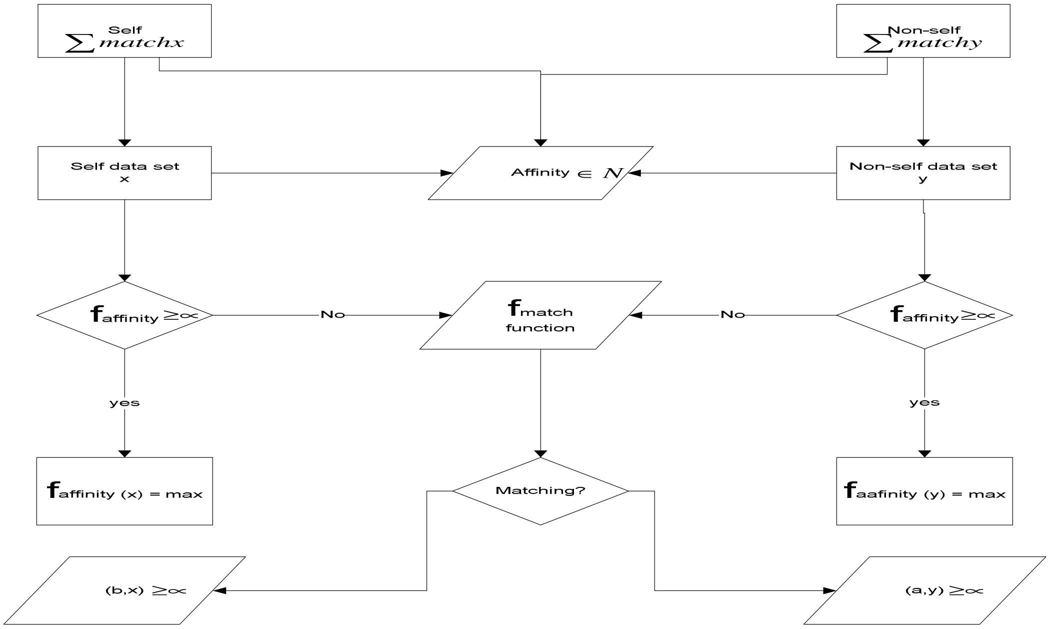

Fig. 2 Matching rule architecture

The matching rules architecture is as presented above where the matching summation of both x and y (self and non-self) is indicated with both data set. An affinity element of N is the ability and number of time to match self and non-self in the dataset. Detail of the matching rule architecture is analyzed in our matching rule model.

6. Experiment and Results

This paper looks at algorithm of spam detection using negative selection based on classification of spam content on a network thereby increasing the accuracy of spam detection with negative selection techniques. The data set used in this technique has 4601 instances in which 39.4% are spasms and each instances has 57 attributes. The data set are loaded and divided in to two classes. The training data set which is meant for training the data and the testing data set which is meant to test the trained data set. The spam e-mail are called spam corpus and are separated from the training data set. The spam corpus is divided in to exemplar and training data set. Let's assume that about 50% of the spam data are exemplars, the exemplar is initialized from random sample of trained data set.

The distance between the exemplar classes and training data sets are calculated by Euclidean distance formulae which is use for data classification. The minimum distance is selected to determine the class of exemplar for classification. The Euclidean formula for continuous value is stated as below.

Total Distance (X1,X2) = $\sqrt{\sum_{i=1}^{n}(X1i - X2i)}$

(25)

X1, X2 and I are the class and specified attribute of exemplar training data.

Matching is applied on the chosen classes of exemplar after initialization. Total rate is the bases for selecting the classes. The percentage of classes which are selected from exemplar randomly is the total rate. Token rate is the percentage of bits from each class which are selected for matching; the bits of class are selected base on token rate of matching. Classification performance is estimated by Euclidean distance formulae at matching class in every iteration. The matching class becomes the new class with better performance. This becomes the new detector with better performance. The distance between the optimized exemplar and training data set are finally estimated and the class with minimum distance value is selected for classification. Performance measurements using the proposed technique were presented in the Table below with different threshold for testing data.

Threshold	Accuracy	False Positive	False negative
0.4	82.7997%	29.9291%	0.1105%
0.5	82.9989%	23.5719%	9.2183%
0.6	86.375%	4.904%	27.5255%
0.7	88.9833%	0.5992%	28.6389%
0.8	89.2658%	0.2252%	29.8358%
0.9	89.2993%	0.1176%	30.6849%

Table 1. False positive and false negative

Given the variation of threshold, the Table 2 above shows how false positive and false negative of our model changes. The accuracy increase with increased in threshold from 0.4 to 0.9, accuracy increase from 82.8% to 89.3% which indicate an increase in the accuracy performance and a reduction of false positive from 29.9% to 0.12% and an increase in false negative from 0.11% to 30.7%. This shows that with improved accuracy performance, there is less false positive (incorrect classification of data as spam) while in improved accuracy, false negative shows an increase (incorrect classification of data as non-spam). This result shows the correctness of self (specific antibody) and non-self (non-specific antibody) of our proposed model. The general result of false positive and false negative help with the assessment of the improvement of negative selection detector or self detector.

7. Conclusion

Computer system complexity is fast becoming a worrying issue and as tremendous influence in spam propagation. Antivirus finds it difficult to detect spam these days as it has become invisible in our computer system. In this paper, we present the self and non-self in a way to create efficiency of detector generation through equation. The novelty of this paper is to generate a new self (system) that randomly create antibody, introducing a new self detector method, with respect to self and non-self producing advance antibody. Also self and non-self matching algorithm is also presented. Mathematical model for effective matching of self and non-self for effective detector is been proposed.

References

[1]. Golovko, V., Et Al., (2010) Neural Network And Artificial Immune Systems For Malware And Network Intrusion Detection, J. Koronacki, Et Al., Editors. P. 485-513.

[2]. Wang, C. And Y. Zhao, (2008) A New Fault Detection Method Based On Artificial Immune Systems. Asia-Pacific Journal Of Chemical Engineering, 3(6): P. 706-711.

[3]. Qing, J., Et Al., (2009) AIS Email Classification Method. Springer-Verlag Berlin Heidelberg 2009. Icic 2009, Lnai 5755: P. Pp. 492–499.

[4]. Zhang, P., W. Wang, And Y. Tan, (2010) A Malware Detection Model Based On A Negative Selection Algorithm With Penalty Factor. Science China Information Sciences, 53(12): P. 2461-2471.

[5]. Akio Ishicuroi, Yuji Watanabe It, And Y. Uchikawa (1994) Fault Diagnosis Of Plant System Using Immune Networks. Proceedings Of The Ieee International Conference On Multisensor Fusion And Integration For Intelligent Systems.

[6]. Hunt, J.E. And D.E. Cooke,(1996) Learning Using An Artificial Immune System. Academic Press Limited. 1084–8045/96/020189+24 $18.00/0

[7]. Stephanie Forrest And A.S. Perelson, (1994) Self Nonself Discrimination In Computre. Ieee. 1063-7109/94 503.00 0.

[8]. Xiaomeng Bian And J. Qiu, (2006) Adaptive Clonal Algorithm And Its Application For Optimal Pmu Placement. Ieee. 0-7803-9584-0/06/$20.00o.

[9]. Vincenzo Cutello, Giuseppe Nicosia, And M. Pavone, (2006) Real Coded Clonal Selection Algorithm For Unconstrained Global Optimization Using A Hybrid Inversely Proportional Hyper Mutation Operator. Acm Transactions On Information And System Security, 1595931082/06/0004.

[10]. Maoguo Gong, Et Al., (2007) Improved Real-Valued Clonal Selection Algorithm Based On A Novel Mutation Method. Ieee Institute Of Intelligent Information Processing, Xidian University, Xi'an 710071, China.

[11]. Jacek Da¸Browski And M. Kubale,(2008) Computer Experiments With A Parallel Clonal Selection Algorithm For The Graph Coloring Problem. Ieee, 978-1-4244-1694-3/08/$25.00

[12]. Xianjin Fang, L.L., (2010) An Improved Artificial Immune Approach To Network Intrusion Detection. Ieee, 978-1-4244-5848-6/10/.

[13]. Hu Zhengbing, Zhou Ji, And M. Ping, (2008) A Novel Artificial Immune Network Algorithm. Ieee 0-7695-3090-7/08 $25.00.

[14]. A. J. Graaff And A.P. Engelbrecht, (2007) A Local Network Neighborhood Artificial Immune System For Data Clustering.. Ieee, 1-4244-1340-0/07.

[15]. Chenggong Zhang And Z. Yi, (2007) An Artificial Immune Network Model Applied To Data Clustering And Classification. Springer-Verlag Berlin Heidelberg P. Part Ii, Lncs 4492, Pp. 526–533,.

[16]. Hengjie Li, Xiaohong Hao, And L. Zhang,(2008) A Clonal Selection Algorithm Based Optimal Iterative Learning Control Algorithm. Ieee Proceedings Of The 7th World Congress On Intelligent Control And Automation Chongqing, China.

[17]. John E. Hunt And D.E. Cooke, (1996)Learning Using Ais. Journal Of Network And Computer Applications (189–212).

[18]. Modupe Ayara, Et Al., Negative Selection How To Generate Detector.

[19]. J.R. Al-Enezi, M.F. Abbod , And S. Alsharhan, (2010). Artificial Immune Systems – Models, Algorithms And Applications. Ijrras.

[20]. Esponda, F., S. Forrest, And P. Helman, (2004) A Formal Framework For Positive And Negative Detection Schemes. Ieee Transactions On Systems, Man, And Cybernetics, Part B: Cybernetics, 34(1): P. 357-373.

[21]. Bindu.M.S And Sumam Mary Idicula. (2011) A Hybrid Model For Phrase Chunking Employing Artificial Immunity System And Rule Based Methods. International Journal Of Artificial Intelligence & Applications (IJAIA), Vol.2, No.4

Author

Ismaila Idris completed is master degree in the area of software engineering at University of Ilorin, Nigeria and obtain his Bachelor of Mathematics and Computer Science degree at Federal University of Technology, Minna, Nigeria. His research interest include intelligent software engineering, intelligent software agent, Artificial Intelligent, Data mining.