A Novel Approach to Building an Artificial Immune Perceptron Using Clonal Selection Algorithm

Hossam Meshref
Associate Professor, Computer Science Department
College of Computers and Information Technology
Taif University, Taif, Saudi Arabia.

Abstract
The artificial immune system (AIS) has many features that could be useful in various research applications. For example, it adapts to changes in its surrounding environment, it has a hierarchical organization, its control is distributed, and it has an ability to discriminate between self and non-self entities. Model design within some of these research applications could sometimes get complicated, making it difficult to use it in some classification tasks. It is believed that to enhance the range of applications, in which the AIS could be used, decomposition of the artificial immune network to its primal elements is required. Hence, in this research, the design of the AIP, as a primary constituent of the artificial immune network, is proposed. Since the AIS and the artificial neural systems share a few common concepts, it is reasonable to compare both perceptron models in both networks to judge the performance of the proposed AIP.

ANN are modeled after the biological neural networks that constitute the basic building blocks of the nervous system. The biological neuron consists of the soma, axons, dendrites and the synapses. The biological neural networks consist of many neurons connected in a specific way to learn to perform a certain function. Think of such functions that a human could need in a daily life. Starting from early childhood, that function could be learning how to hold a cup to drink, and to adjust the hand eye coordination mechanism using these networks to perform such tasks. As the network learns to perform a certain function, the whole experience is stored in the synaptic strength between neurons. Some of the well-known original definitions for the ANN is:

"An artificial neural network is a massively parallel distributed processor that has a natural propensity for storing experimental knowledge and making it available for use"[1]. The focus in this research will be on the ANP model as a model to compare its performance with the designed AIP model. The theory and algorithms behind the ANP will be discussed shortly in the design section. Having definedANN, it is logical to define the AIS as well. According to Dasgupta and Nino, immunity could be regarded as:

“The condition in which an organism can resist diseases, more specifically infectious diseases. However, a broader definition of immunity is the reaction to foreign substances, pathogens, which includes primary and secondary immune responses.”[2]. The biological immune system consists mainly of lymphocytes that have two major types, T-cells and B-cells. B-cells are responsible for humoral immunity that secretes antibodies. On the other hand, T-cells are responsible for cell mediated immunity. Each B-Cell has a unique structure that produces suitable antibodies in response to invaders of the system. That type of response is called innate immunity and eventually results in antibody-antigen relations to be stored in case the host encounters the same invader again. In that case, the immune response in expected to be faster given that the network has seen it before, i.e. learned how to deal with it [3].

Keyword: Artificial Immune Systems; Artificial Neural Networks; Artificial Neural Perceptron; Artificial Immune Perceptron

I. INTRODUCTION
A few biologically inspired techniques have been implemented in different areas of research. For example, techniques based on neural networks and genetic algorithms were deployed in many research projects. However, there is a new area of biology that has been utilized recently by the computational intelligence community: the artificial immune systems. There were many research efforts that utilized the artificial immune networks to model their applications. However, the models they built were considerably complicated. That led researchers in this field to be restricted to a few areas that are based on self/non-self identification feature in the immune system such as computer security applications.

It is believed that in order to enhance the range of applications in which the artificial immune network could be used, decomposition of the artificial network to its primal elements is required. Hence, in this research, the design of the AIP, as a primary constituent of the artificial immune network, is proposed. Since the AIS and the artificial neural systems share a few common concepts, it is reasonable to compare both perceptron models in both networks to judge the performance of the proposed AIP.

"An artificial neural network is a massively parallel distributed processor that has a natural propensity for storing experimental knowledge and making it available for use"[1]. The focus in this research will be on the ANP model as a model to compare its performance with the designed AIP model. The theory and algorithms behind the ANP will be discussed shortly in the design section. Having definedANN, it is logical to define the AIS as well. According to Dasgupta and Nino, immunity could be regarded as:

“The condition in which an organism can resist diseases, more specifically infectious diseases. However, a broader definition of immunity is the reaction to foreign substances, pathogens, which includes primary and secondary immune responses.”[2]. The biological immune system consists mainly of lymphocytes that have two major types, T-cells and B-cells. B-cells are responsible for humoral immunity that secretes antibodies. On the other hand, T-cells are responsible for cell mediated immunity. Each B-Cell has a unique structure that produces suitable antibodies in response to invaders of the system. That type of response is called innate immunity and eventually results in antibody-antigen relations to be stored in case the host encounters the same invader again. In that case, the immune response in expected to be faster given that the network has seen it before, i.e. learned how to deal with it [3].
II. BACKGROUND

Surveying the literature, it was found that a few researchers worked on a line of research trying to use hybrid system of artificial immune systems and other systems, e.g., Hopfield neural networks, fuzzy logic, and genetic algorithms [4,5,6,7]. For example, Kuo, Tseng, and Chen proposed a novel approach to solve the optimal problems of multimodal function with high dimensions. They started with the AIS algorithms to get initial solutions for the problem and fed that to the K-mean clustering algorithm. The cluster centers are then used as initial values for the Hopfield network to obtain all the minimum extreme values. There results proved that the new hybrid algorithm is faster and more accurate than trying solely traditional artificial immune algorithms. On the other hand, Yue et al had a different another approach that enhanced the performance of the back-propagation network in a power system short-term load forecasting. Their modified AIS algorithm helped in optimizing weight and threshold values for their designed back propagation neural network. That hybrid designed model findings showed higher forecasting accuracy compared with the artificial back propagation neural networks model alone.

Aside from the hybrid systems research attempts, other researchers surveyed a lot of research work on security applications [8,9,10]. They commented on some of the challenges that the AIS may encounter in such security applications such as the dependence of normalcy of network traffic time. What should be considered normal or abnormal based on the network measurement time: day, night, or a weekend. These challenges created an urgent need for the design of an analysis model that can handle multi-variable settings and large volumes of data rather than just tackle each parameter alone. These systems were not always based solely on the artificial immune system, but sometimes complement its function to detect abnormalities with other bio-inspired techniques such as artificial neural networks, genetic algorithms, swarm intelligence, Swarm Intelligence, and Learning Classifier Systems. The combination of the aforementioned algorithms, in spam filtering applications for example, could enhance the classifier capability as it handles false positive and detection rates.

As a continuation in the research direction of anomaly detection, the area of fault detection and recovery had a considerable amount of research as well. For example, a few researchers focused mainly on the AIS techniques and its biological features that could be modeled mathematically to tackle fault detection and recovery challenges [11]. In that research survey, researches work utilizes some features of the artificial immune system. These features base on its ability to perform positive and negative selection for example. The surveys describe one-signal fault detection approach, where positive detection as well as negative selection artificial immune algorithms are used. another approach is the two-signal fault detection approach, where techniques such as danger-model based as well as NK-cell based were used.

A similar research work has been done in a chemical fault detection and diagnosis application using the pattern recognition ability of the AIS [12]. The designed model was able to take corrective actions as well. The researchers utilized the AIS model features but modified the affinity measure mechanism by using principle components analysis. They were able to model adaptive immunity and continuously build knowledge about the undergoing chemical processes. It is evident in the aforementioned fault detection and recovery application that the AIS features can solely solve that problem without hybridization with other similar biologically inspired models.

Most of the aforementioned research made considerable efforts in shedding some light on the importance of ANN as well as the artificial immune networks, or a hybrid system of both. The goal of those hybrid systems was to get the best advantages of the combined systems, but the resultant systems were relatively complicated. These complications come with a cost that urge researchers to seek simpler solutions that do not compromise efficiency and at the same time adhere to simplicity. However, integrating both network paradigms is not the goal of this proposed research, and neither is the replacement of ANN by a more robust artificial immune network or vice versa. What is being approached here is a novel attempt in which the complexity of the artificial immune network is being decomposed to its primal element: the AIP. The proposed artificial immune network design model is elaborated to encompass a multi-layer AIP, and to test network performance with different experiments to test the worthiness of the new proposed model. A comparative analysis with the corresponding ANP is used in these experiments as a reference for a well known designed model in pattern recognition. These proposed models explore of the abundance of opportunities that is expected to lay a foundation of an array of research projects that will benefit from the findings of the proposed research. One major advantage is to produce a simple, yet a robust, model for an artificial immune network that could be useful in many applications. Moreover, producing a simplified model of AIN is expected to encourage researchers in the field to deploy the proposed designed model in their applications as well as enhance it to develop advanced practical models.

The next section, section two shed light on the artificial neural network perceptron. Section three describes the proposed AIP structure. Research results are discussed in section four, followed by the conclusion along with the future work research directions for interested researchers in section five. The designed multi-layer perceptron model is expected to be useful in similar applications that focus on improving the learning process within the artificial immune network.
III. THE ANP

The first concept proposed in ANN is transforming the biological neuron into the artificial neuron by building the Perceptron, see figure 1(a).

![Schematic model of a biological neuron](image)

**Figure 1 (a) THE BIOLOGICAL NEURON**

In figure 1 (b), a set of n synapses is associated to the inputs. Each of them is characterized by a weight. A signal at the input is multiplied by its corresponding weight, and all the weighted input signals are summed. Thus, a linear combination of the input signals is obtained. A "free weight", or bias, is added to this linear combination and this forms a weighted sum, see equation 1.

\[ Z = \sum_{i=1}^{n} x_i w_i + x_0 w_0 \]  
\[ y = \phi(Z) \]

**Equation 1**

**Equation 2**

A nonlinear activation function \( \phi \) is applied to the weighted sum, and eventually the value of the activation function is the neuron's output, see equation 2.

**Figure 2**: Activation Functions

(A) Logistic Function  (B) Hyperbolic Tangent Function

During the training of the perceptron, the weight parameters are adjusted as the actual target is compared with the target output. If the actual output is similar to the target output, no change in weight is required. Otherwise, weight update is required according to the following equation, equation 3:

\[ W_{k+1} = W_k + \lambda (o_i - t_i) x_{ij} \]  

**Equation 3**

Where \( \lambda \) is the learning rate, and \( W \) is the weight connecting neuron \( i \) and neuron \( j \). In equation 3, it is evident that the required weight change is proportional to the error \( (o_i - t_i) \). The learning process of the perceptron is governed by the following algorithm:

**The Perceptron Learning Algorithm**

1. Initialize (weights);
2. For (t=0; t<Sizeof TrainingFile; t++)
   - Compute(Y_net);
   - ApplyThreshold(Y_net);
   - Compute(Y_out);
   - If (Yout!=TargetValue)
     - UpdateWeights();
     - Continue;
   - Else
     - Break;
3. Having explained the basic concept behind the ANP, which is its basic building block, we can start analyzing the corresponding version in the artificial immune network. As the neuron is the basic building block of the nervous system, the B-Cell is the main building block of the Idiotopic immune network. In the following section, the concept of the AIP will be discussed and gradually built upon using concepts drawn from the biological immune system.

IV. THE PROPOSED AIP

According to Decastro and Timmis and Harmer et al, to apply an immunity based model in a specific domain, we should follow a series of problem solving stages to find a solution for a particular problem [13, 14]. In this research design, attention is
focused on the Biological B-cell and its corresponding mathematical model that could be utilized to build the AIP, see figure 3. Within the B-Cell, two stages are represented: affinity measurement and affinity maturation, see figure 4. In affinity measurement the process of Ag-Ab matching is done, and matching pairs are produced: \((P, Ep)\). As for the affinity maturation stage, only the matching pairs with high affinity (greater than a certain threshold) are going to be cloned and circulated in the artificial immune network.

**FIGURE 3 COMPLEMENTARITY BETWEEN THE BINDING REGION OF A RECEPTOR AND AN ANTGEN EPITOPE [15]**

Starting with the affinity measurement stage, the immune system antigen recognition is based on the concept of complementarity between the binding region of the receptor \(P\) and the Ag epitope \(Ep\). When complementarity is satisfied, then the B-Cell knows that this Ab locks to that Ag. This process is called *affinity*. There are many equations to measure the degree of affinity. Among which we will focus on the bit-wise affinity, where the string matching rule depends on the representation scheme and type of data. Therefore, that rule the hamming distance between two strings \((P,Ep)\) could defined using equation 4:

\[
h(P, Ep) = \sum_{i=1}^{N} P \oplus Ep \quad \text{eqn. (4)}
\]

\(h(P, Ep)\) is defined as the number of different bits between the two strings \(P\) and \(Ep\). For example, if \(P=[1 \ 0 \ 0 \ 1 ]\) and \(Ep=[\ 0 \ 1 \ 0 \ 0 ]\) then we expect \(h(P, Ep)\) to be equal to 4, which is the length of the string \(P\) or \(Ep\). That case is considered to have full complementarity where all the corresponding bits differ and as a result we have: \(h(P, Ep) = N,\) where \(N\) is the string length. However, as shown in equation 5, the calculated affinity \(\mathcal{A}(P,Ep)\) in general must exceed \(\theta\) (threshold) and not necessarily satisfy full complementarity, and \(a_i\) in that case measures the complementarity level.

\[
\mathcal{A}(P,Ep) = \begin{cases} 
0, & h(P, Ep) < \theta \\
1, & h(P, Ep) \geq \theta 
\end{cases} \quad \text{eqn. (5)}
\]

Figure 5 illustrates the interaction that is taking place within the affinity measure stage starting from measuring the hamming distance between each \(P\) and \(Ep\) and ending up with deciding on the affinity levels.

All the aforementioned AIP design is based on an artificial immune model that considers the interaction between \(P\) and \(Ep\), see figure 6. The affinity maturation stage, in general, is expected to produce clones on high or low affinity. In this research, the clonal selection algorithm is going to be used to control the learning process with the AIP, see figure 7.
V. EXPERIMENTAL TESTS AND DISCUSSION OF RESULTS

Consider a simple ANP that has a single binary decision output unit. There are two input images that the perceptron needs to recognize: a 3x4 image, $S_1$, representing number 1, and another 3x4 image, $S_2$, representing number 7, see figure 8. Pixels in images $S_1$ and $S_2$ are considered bipolar, which means it is either 1 or -1.

Figure 8 The Two Input Patterns

A. Character Recognition Using the ANP

Let us assume that there are a set of vectors in $\mathbb{R}^n$ where they could be classified into one of two classes $S_1$ and $S_2$. Each vector is in the format of $(x_1, x_2, ..., x_n)$, $n=12$. The output of the ANP is a linear combination of its weighted inputs, see figure 9, next page. The goal of the Perceptron algorithm is to find the values of those weights $w_1, w_2, ..., w_n$ that will enable the perceptron to differentiate between the two classes $S_1$ and $S_2$. The training of this proposed perceptron model is considered supervised. This means that the perceptron is expected to produce an output equals to 0 if class $S_1$ is recognized and produce an output equals to 1 if class $S_2$ is recognized. The ANP was trained for different epochs, and the recognition error values were recorded. Until epoch 50, the error was 0.9 and even after 500 iterations, where the error was almost 0.889. Figure 10, next page, shows only 50 iterations, because minor improvement was observed in the successive iterations.

Figure 9 The Basic ANP Model

Figure 10 Training of the Basic ANP Model

B. Character Recognition Using the AIP

Based on the previous description of the AIS components we can describe their interaction. The clonal selection algorithm is one of the methods to model such interactions of the AIS with the external environments, or in other word, invading antigens.

Figure 11 Initial Population Sample

Similar to the ANP algorithm, the clonal selection algorithm could be used to distinguish between two classes. In this simulation, One AIP unit is used to be compared to the basic ANP unit. In figure 11, $P_1$, $P_2$, ..., $P_6$ are the initial population samples, and they are considered to be the antibodies, or antigen receptors, while $S_1$ and $S_2$ are considered antigens. It is assumed that the elements of $S$ and $P$ are represented in the same shape-space $S'$. In the first iteration, after running the algorithm, we calculated the affinity for each antigen, $M_1$, with elements of $P$: $M_1 = \begin{bmatrix} 6 & 2 & 1 & 6 & 6 \\ 2 & 6 & 5 & 6 & 2 \end{bmatrix}$

As mentioned earlier, the affinity values range between 0 and 12. According to the algorithm, the second step is to perform clonal selection and expansion. The highest $n_1$ affinity elements of $P$ are chosen ($n_1=3$ in this experiment), and clones of these elements are generated in proportion to their affinity with the antigen. As the affinity increases, the number of copies increases. Next, Mutation occurs in proportion to the individual’s affinity as well. As the affinity increases, the mutation rate decreases.

After Mutation, the mutated elements are added to the original population $P$. Then, the best individual is reselected to be kept as memory. As part of the meta-dynamics process, the lowest $n_2$ affinity elements of $P$ are chosen ($n_2=2$ in this experiment). Then, they are replaced by a randomly generated new individuals. The whole process is repeated until a stopping criterion is met. In this experiment, the stopping criterion is met, when any affinity matching reaches 12. Figure 12 shows the affinity maturation for $S_2$, where a mutated version of $P_2$ at iteration 40 was able to recognize its pattern.

Figure 12 Training of the Single AIP
C. Character Recognition Using the single-layer ANP

In the previous ANN experiment, the error showed no improvement less than 0.8, even after 500 iterations. It was imminent to increase the number of perceptrons per layer initially, in a single layer, to be able to judge on the network performance and eventually compare the findings with the AIP performance. A 7-1 single-layer perceptron was used to that end, see figure 13. Not too much improvement has been observed compared to the previous single neuron experiment. After 50 iterations, a 0.80963 error was achieved and no noticeable progress took place even after 500 iterations; the error became 0.80492, see figure 14.

D. Character Recognition Using the single-layer AIP

The previous experiment in section B was repeated using a single-layer of AIPs. Seven AIP’s were used in a single-layer, and the simulation learning process was controlled by the clonal selection algorithm, see figure 15.

E. Character Recognition Using the multi-layer ANP

In the depicted design, we have seven AIP’s in one layer, which is considered the output layer. The AIP clonal selection algorithm is run for each AIP element in the layer, and the output of the winning element is chosen. The winning element is the one with the highest affinity. For each AIP we used a different input P vector ($p_1, p_2, p_3, p_4, p_5$).

After 14 iterations, for the $S_1$ pattern, $P_2$ vector was able to achieve full maturation before other recognizing vectors, which is an almost Zero Error recognition rate. Since the stopping criterion has been met, there was no need to further train the network using the other $P$ vectors. At this point, the final mutated $P_2$ is the considered as the best recognizing vector (antibody). Then, eventually is reselected to be kept as memory, see figure 16.

To be able to accurately judge the recognition ability of the AIP, it had to be compared with a full multi-layer artificial neural perceptron model, and eventually compare the findings with the AIP performance. A 7-5-1 multi-layer ANP was used for that end, see figure 17. After training the network, a considerable improvement has been observed compared to the single-layer ANP. The recognition error reached 0.00982 at the 20th iteration, 0.000728 at the 200th iteration, and eventually 0.000191 at the 500th iteration, which is almost Zero, see figure 18.
F. Character Recognition Using the AIP

experiment done at section D, stage-1, was continued to investigate the overall network performance by adding a second layer of 5 AIPs, as we finished stage-1, and the winner AIP was identified, see figure 19. That vector was less mutated and more cloned, while other AIP vectors, with less affinity were more mutated and then less cloned. From that pool, the inputs of the second artificial immune layer were chosen. It worth mentioning that that same S1 and S2 from the original input layer are used as inputs to that second (output layer) as well. The clonal selection algorithm was run for each element in the output layer, and the output of the winning element was chosen. At this point, P2 is the considered as the best recognizing vector (antibody). Then, eventually is reselected to be kept as memory, see figure 20. Similar process was used to find the best recognizing vector for S2, and it was found that the mutated version of P3 can meet the condition for the stopping criterion after 9 iterations, see figure 21.

VI. CONCLUSION AND FUTURE WORK

In this paper, simple AIP models are presented, and compared with the corresponding parallel models of the ANP models. Within the AIP models, the error rate decreased by 10% between the 1-1 and the 7-1 network structures. However, moving from the 7-1 to the 7-5-1 structure, the error rate noticeably improved from 0.8 to almost 0.01. That result is expected based on the 3-layer artificial neural network recognition capability. On the other hand, within the AIP model structures, the recognition performance increased by 67% between the 1-1 and the 7-1 network structures, while recognizing pattern S1. As for pattern S2, the recognition performance increased by 50% between the 1-1 and the 7-1 network structures. Moving to the multi-layer AIP, the recognition performance increased by 28.6% between the 7-1 and the 7-5-1 network structures for pattern S1. As for pattern S2, the recognition performance increased by 55% between the 7-1 and the 7-5-1 network structures.

The previous findings reflect that as the number of units in both networks increases, the recognition ability increases, and the error rate is reduced. However, it is imperative to judge on the performance of the AIP network by comparing it with the artificial neural network to evaluate the performance of the novel design. For the 1-1 network design, the AIP design outperformed the AIP significantly as it converged after 42 iterations for S1 and 40 iterations for S2 compared to an error rate of 0.9 even after 500 iterations. For the 7-1 design, as well, the artificial immune network model continued to outperform the artificial neural network model as it converges after 14 iterations for S1 and 20 iterations for S2 compared to an error rate of 0.8 even at 500 iterations. As both networks have
more complex structures, 7-5-1, the performance gap was decreased, but the ANP network model still outperformed the ANP network model. The artificial immune network model converged after 10 iterations for S, and 9 iterations for S2 compared to an artificial neural network convergence after 20 iterations. Particularly, it was noticed that the AIP network-based model simulation examples show a better convergence and recognition abilities compared with the corresponding ANP network-based model examples.

As for the future research directions, the proposed artificial immune network could be improved by designing a reverse error calculation pass and eventually compare the overall AIP network-based model findings with a back-propagation artificial neural network model. That research direction could open new avenues based on the new model. On the other hand, other artificial immune network training algorithms, such as the negative selection algorithm, could be also investigated to study its effect on the pattern recognition process, and perhaps eventually could be compared with the clonal selection algorithm.

REFERENCES