International Journal of Pattern Recognition
and Artificial Intelligence
Vol. 25, No. 8 (2011) 1275–1295

DOI: 10.1142/S0218001411009123





# SPAM DETECTION USING GENETIC ASSISTED ARTIFICIAL IMMUNE SYSTEM

RAED ABU ZITAR
*School of Engineering and Computing Sciences*
*New York Institute of Technology, Amman, Jordan*
*rzitar@nyit.edu*

ADEL HAMDAN MOHAMMAD
*Computer Information System Department*
*Applied Science University, Amman, Jordan*
*A_hamdan@asu.edu.jo*
*Adel_hamdan@yahoo.com*

This work presents a novel system based on artificial immune system for spam detection. A relatively new machine learning method inspired by the human immune system called Artificial Immune System (AIS) has been emerging recently. This method is currently undergoing intense investigation and demonstration. Core modifications were applied on the standard AIS with the aid of the Genetic Algorithm (GA). SpamAssassin corpus is used in all our simulations. Spam is a serious universal problem which causes problems for almost all computer users. This issue affects not only normal users of the internet, but also causes problems for companies and organizations due to expensive costs in lost productivity, wasting users' time and network bandwidth. Many studies on spam indicate that it costs organizations billions of dollars annually. We introduce a GA assisted AIS in spam detection, and compare between two methods. Encouraging results were achieved when comparing to commercially available anti-spam software.

*Keywords*: Artificial immune system; genetic algorithm; spam detection.

## 1. Introduction

E-mails are one of the most important forms of communication; e-mails are simple, effective, and a cheap type of communication for almost all computer users. This simplicity and cheapness are prone to a lot of threats. One of the most significant of them is spam; spam e-mails are a problem that almost every e-mail user suffers from. The word "spam" usually denotes a particular brand of luncheon meat, but in recent times, spam is used to represent a variety of junk, unwanted e-mails. It is now possible to send thousands of unsolicited messages to thousands of users all over the world at approximately no cost. As a result, it is becoming common for all users worldwide to receive hundreds of spam messages daily.[6,9]

There are several approaches which try to stop or reduce the huge amount of spam which target individuals. These approaches include legislative measures such as worldwide antispam laws. Other techniques are known as Origin-Based filters which are based on using network information and IP addresses in order to detect whether a message is a spam or not.[6,9] The most common techniques are filtering techniques, attempting to identify whether a message is spam or not based on the content and other characteristics of the message.[18] In spite of the large number of methods and techniques available to combat spam, the volumes of spam on the internet are still rising.[1,4,20,22,26]

This work presents a new solution for spam inspired by Artificial Immune System model (AIS). With the help of Genetic Algorithm (GA) a lot of modifications on standard artificial immune system are sought in order to make it work more efficiently. This work also includes a comparison study between genetic optimized spam detection using AIS and standard AIS for spam detection.

### 1.1. *Problem statement*

Nowadays, spam has the potential ability to become a very serious problem for the internet community, antispam vendors offer a wide array of products designed to help us keep spam out. They are implemented in various ways (software, hardware), several techniques (content, rule-based) and at various levels (server and user). The introduction of new technologies, such as Bayesian filtering, Support Vector Machines (SVM), Artificial Neural Network (ANN), Artificial Immune system (AIS), etc. can improve the accuracy of filters.[4] The implementation of machine learning algorithms is likely to be very important in the continuous fight against spam.[13,16,20,23,26]

AIS is a new paradigm that can be classified as a knowledge based system technique which implements machine learning and develops its own library or knowledge base. In this paper, Artificial Immune System (AIS) is used in spam detection and Genetic Algorithm (GA) in optimizing the antispam Artificial Immune system. Few previous studies have used AIS in spam detection and the subject still needs more investigation. No previous work has tried to optimize the large number of parameters of the AIS, especially in an important application such as spam detection. In this paper, a lot of parameters are modified or optimized to get more accurate results, the advantages of the all-purpose optimization embedded in the genetic algorithm (GA) will be used in order to achieve an optimum AIS. Several analysis and comparisons will be made. A valuable antispam software product is expected to be developed as an outcome of this paper. Research outline is as in Fig. 1.

The contribution of this work can be summarized as follows:

(1) Demonstrating the use of AIS on spam detection. This subject is still new and requires more applications, verification and testing.
(2) Using (GA) to optimize AIS spam detection in:
    (a) Determining when to perform culling (replacing old lymphocytes with new ones).

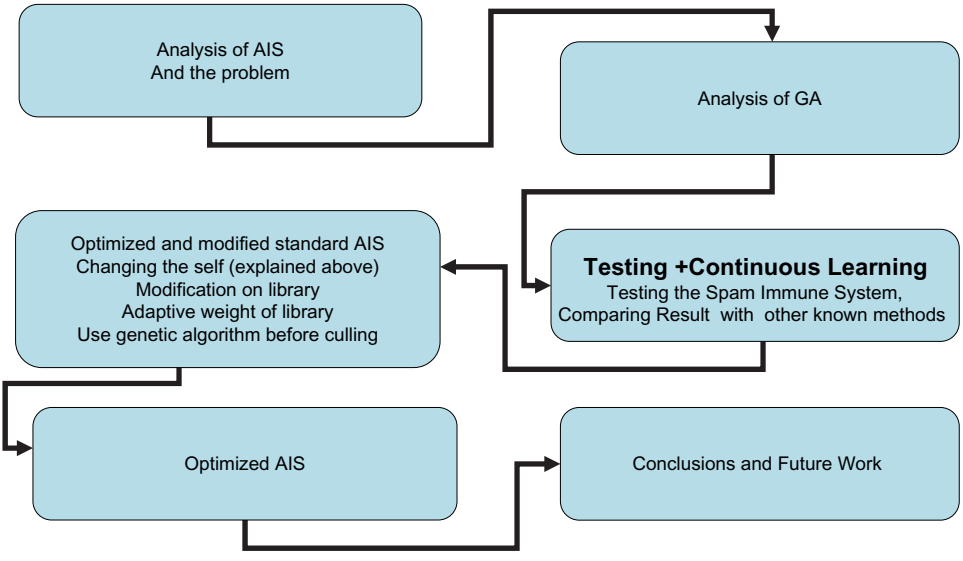


Fig. 1. Research outline in general.

(b) Determining when to check if the self (legitimate) is changed (to fit the new interest of users).

(3) Developing a new approach for learning in AIS that allow new adaptive immune system (lymphocyte) to take place instead of useless lymphocyte in innate immune system.

(4) Applying the different techniques in spam detection and comparing and analyzing the results.

## 2. Immune System Fundamentals

The immune system is very complex and its complexity varies according to its characteristics. For example, some plants have protective spines to provide protection from predators that attack them. Animals have bones (vertebrates) which contain a developed and a highly effective and complex immune system. It consists of a vast array of cells, molecules and organs that work together to maintain and keep life.[2,7,10,19] The focus here will be on the immune system of vertebrates, more specifically of humans. This is because of its interesting features, characteristics from a biological and computational perspective. Extensive knowledge is available on its implementation and its broad applicability in the design of AIS.[3,5,14,25]

The immune system performs several functions. One of its main functions is that the immune system together with other bodily systems maintain a constant state of essential functions, named as homeostasis. One of its most amazing roles however is

the protection of the organism against any foreign attack of disease which may cause agents, called pathogens, and the exclusion of malfunctioning cells.[8,12,17,21]

## 3. How Genetic Optimized Spam Detection Uses AIS Works

In the following sections, the main components of spam detection system are explained which will be modified to enhance the system. Also, our proposed algorithms will be shown.

### 3.1. *Genetic optimized spam detection using AIS*

GA is an optimization algorithm which can be used in different applications, in standard AIS there are many parameters defined by the user. One of the most important parameters is culling. This parameter is defined by the user in standard AIS. In this work, GA is used to determine the culling time. Also, GA is used in determining the Rebuild time to solve the problem of users' interests which do not remain the same over time.[15,24] Antispam solutions have to increase the frequency of the updates and also to develop more heuristics in less time. The need for an automatic process that would quickly learn the characteristics of the new spam without affecting the accuracy of detection on less recent spam has become vital.

AIS parameter of AIS that is a subject for modification and optimization are the following:

**(1) The Self (legitimate)**

The interest of any person is not stable because circumstances always change. The problem found in the standard immune system (AIS) is the self (Legitimate) which changes over time. This means that the message content and characteristics that any person would like to receive is changed over time. For example, a healthy person might have no interest in any message about medicine. However, if that person is diagnosed with any disease, he may start to receive them. Also any person who has no interest in sports, does not wish to receive any sports messages, however, if that man becomes overweight he may become interested in those messages to become fit. This means that the system must be adapted to this change. This work tries to solve this problem with the help of GA, in this work the system is rebuilt periodically, so the system has the chance to relearn. This means that the system must accept some types of messages which could be recognized as spam and if the user still recognizes this type of message as spam, as a result there is no change on the self. Moreover, the Genetic Algorithm is used in producing guided random rebuild time instead of a fixed period.[11]

**(2) Library**

The standard Artificial Immune System uses a library that does not change over time. But as the system sees more spam messages, the system must have the ability to gather information from these spam messages that could be used to create a new

useful gene fragment. By adding this extra ability which is important to the system to adapt to the gene library, it would be possible to make the system adapt to new messages which are not matched by any current gene fragment. In this work, information is collected from spam messages. Frequent patterns are used to create new antibody or gene library.

**(3) Adaptive weight of gene library**

In standard Artificial Immune System, each gene fragment has an equal chance of being selected, this means that if there is a fragment rarely used, this gene fragment will have an equal chance with any gene that is frequently used. This shows that the system is not adapted. Weights are adapted for each gene fragment based on the previous run.

**(4) Lifecycle of Lymphocytes**

The main point here is to give the lymphocytes a chance to match and to assign a weight but this time must be bounded so that the system does not waste time on lymphocyte that do not match any message "un-useful lymphocyte". In standard AIS after some time has passed (the time interval is chosen by the user), the lymphocytes are aged and may be killed. This amount of time is a parameter of the system, defined by the user. The best choice of an update interval will depend upon the number of

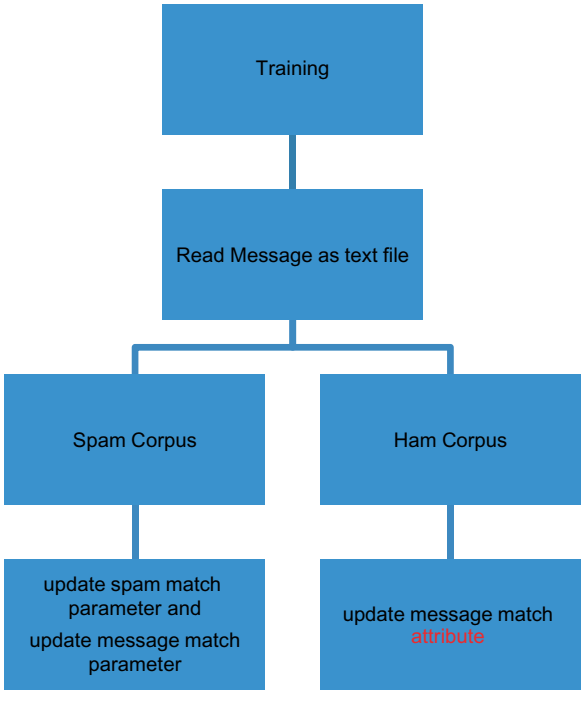


Fig. 2. Training outline.

e-mails received by the users. In this paper, it is defined by using the Genetic Algorithm. This is done in a fashion similar to the approach used with the self.

The modifications on algorithms will be explained in detail in the next section.

### 3.2. *The proposed model*

In genetic optimized spam detection using AIS, there are several major steps:

- Training:
  The aim of training (Fig. 2) is to build a library and from this library, the best lymphocytes will be chosen to fight against spam. In summary, the steps of training are the following:
  - The system reads each message as a text file and then it is parsed to identify each header information (such as "From", "Received", "Subject" and "To").

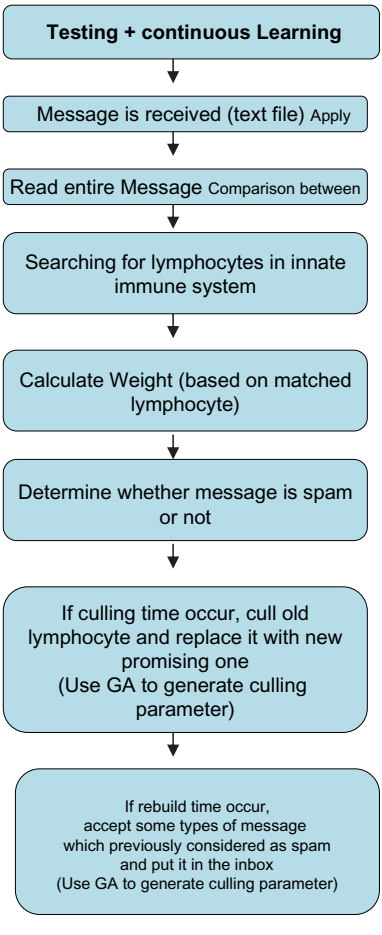


Fig. 3. Testing + Continuous Learning outline (genetic optimized AIS).

- Any lmphocyte which exceeds 20 characters or less than 3 characters is excluded.
- All accepted lymphocytes are added to the library.
- Modification is done on (Spam_matched, Msg_matched) parameters.
- Perform extra cleaning (if necessary) on library to refine it.

- Testing:
  In testing (Fig. 3) there are several steps:
  - When a message is received, the system compiles it as a text file. Then, the system will look in the innate immune system to search for any matched lymphocyte. Then the system calculates the score for each received message to determine if the message is spam or not.
  - Depending on the score, the system will decide if the message is spam or not (if score is greater than threshold then message is spam).



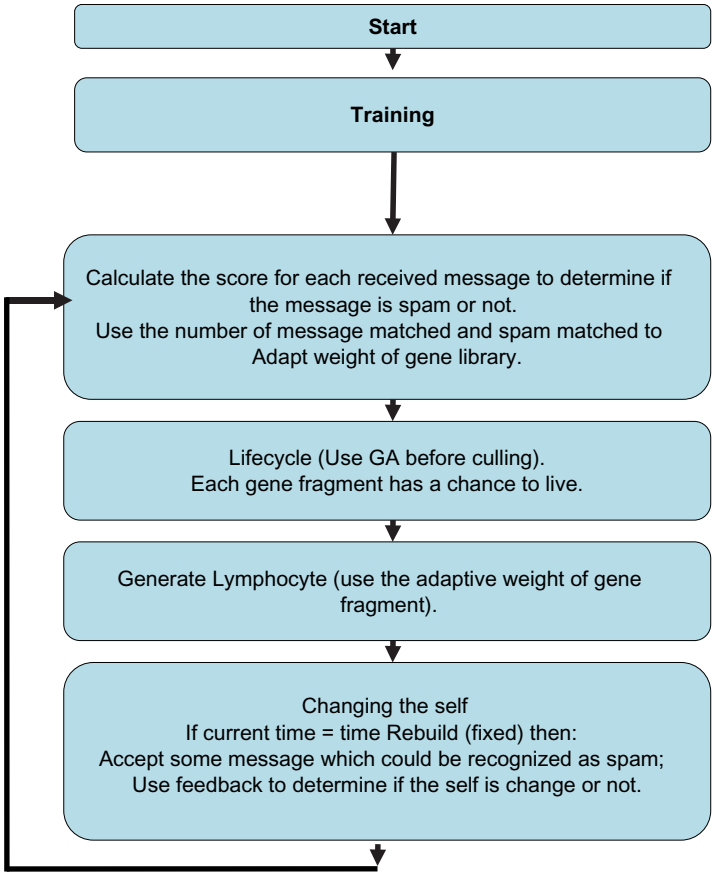


Fig. 4. Genetic optimized AIS process flow (training + testing).

- If the message is spam, then the system will add each new lymphocyte to the adaptive immune system (library) to be used in future (learning).
- The system will use GA to generate a culling parameter. When culling occurs, useless lymphocytes will be deleted and replaced by new promising lymphocytes from the library.
- Also, the system will use GA to generate a rebulid time parameter. When the rebuild time occurs, the system will accept a new type of message which is considered as spam and will put them in the inbox. Depending on the users' feedback, the system will determine if the self is changed or not.

GA in genetic optimized spam detection using AIS is called twice (Fig. 4); firstly, to cull old lymphocytes (useless lymphocytes replaced by new promising ones). Secondly, to check whether there are any new interests for users (changing the self) in a similar manner. This can be done taking into consideration that the system uses different domain ranges for each one.

**Algorithm 6: Genetic optimized spam immune system using genetic algorithm.**

Step 1: Require: Update_Interval for culling old lymphocyte: (Update based on time or based on number of message, in both situations, it is generated based on Genetic Algorithm.
Step 2: Innate_Immune_System ⬅ Ø {Initialize innate immune system (list) of lymphocytes to be empty}
Adaptive_Immune_System ⬅ Ø {Initialize innate immune system (list) of lymphocytes to be empty}
Step 3: Update: If update based on time used.
Update ⬅ current time + update time using GA
Or
If update based on number of message.
Update ⬅ number of message using GA
Start Training (Algorithm 2)
**While** Optimized_Immune_ System is running **do**
    **if** message is received **then**
    Start Application (Algorithm 3)
    **end if**
        **if** current time > update time **then**
        Or
        **if** number of message received > number of message for update) **then**
        Start Learning (Algorithm 4)
        **end if**
**end while**

**Algorithm 7: Training.**

```
Message ⬅ spam or non spam message. (Training corpus)
Innate_Immune_System ⬅ table (may be empty)
Spam corpus
For each lymphocyte in the spam message corpus do
    If lymphocyte is already exist in Innate_Immune_System then
    lymphocyte.msg_matched ⬅ lymphocyte.msg_matched + 1
    lymphocyte.spam_matched ⬅ lymphocyte.spam_matched + spam_increment
    else
    Add lymphocyte to Innate_Immune_System
    lymphocyte.msg_matched ⬅ lymphocyte.msg_matched + 1
    lymphocyte.spam_matched ⬅ lymphocyte.spam_matched + spam_increment
    end if
end for
Ham corpus
For each lymphocyte in the spam message corpus do
    If lymphocyte is already exist in Innate_Immune_System then
    lymphocyte.msg_matched ⬅ lymphocyte.msg_matched + 1
    end if
end for
End
```

### 3.3. *Components of the genetic optimized spam detection AIS*

#### 3.3.1. *The library*

In genetic optimized spam detection AIS training phase is used to create the library which will then be used to generate lymphocytes. The library is created based on training phase results. This library contains thousands of lymphocytes but the majority of these lymphocytes are useless. There are thousands of lymphocytes which appear just one time after completing the training phase. So, in order to decrease the huge number of lymphocytes in the library, all lymphocytes with spam_matched = 1 and Message_matched = 1 are deleted. Useful lymphocytes are those which have matched to a large number of spam messages.

In training phase, each e-mail message is compiled as a text file, and then parsed to identify each header information (such as "From", "Received", "Subject", or "To") to distinguish them from the body of the message. Every substring within the subject header and the message body that was delimited by white space were considered to be a token (lymphocyte).

#### 3.3.2. *Lymphocytes generation*

For the purpose of using biological immune system in spam detection, the term "digital lymphocyte" refers to:

- Digital antibody.
- Weighting information.

**Algorithm 8: Application.**

Step 1: Innate_Immune_System ⇦ the list of Anti_spam lymphocyte
Step 2: Adaptive_Immune_System: Empty Table
Step 3: Message ⇦ a message to be known whether it is spam or ham
Threshold ⇦ a cutoff point valued between 0 and 1 inclusive; anything with a higher score than this is spam {chosen by user}.
Require: increment ⇦ increment used to update lymphocytes
total_spam_matched ⇦ 0
total_msg_matched ⇦ 0
**for** each lymphocyte in Innate_Immune_System **do**
    **if** lymphocyte.antibody matches message **then**
    total_spam_matched ⇦ total_spam_matched + lymphocyte.spam_matched
    total_msg_matched ⇦ total_msg_matched + lymphocyte.msg_matched
    lymphocyte.msg_matched ⇦ lymphocyte.msg_matched + 1
**end for**

Score ⇦ total_spam_matched/total_msg_matched
**if** score > threshold **then**
    Message is spam
    lymphocyte.spam_matched ⇦ lymphocyte.spam_matched + increment
        **if** lymphocyte.antibody does not exist in Innate_Immune_System **then**
        Add lymphocyte.antibody to Adaptive_Immune_System
        (This is to represent continuous learning)
        **end if**
**else**
Message is not spam
**end if**

**End**

**Algorithm 9: (Learning) Cull old lymphocytes and generate new lymphocyte.**

Culling can happen based on
Criteria 1: Number of message calculated using Genetic algorithm
Or
Criteria 2: Update interval calculated using Genetic algorithm
**if** criteria happened **then**
Merge Adaptive_Immune_System with Innate_Immune_System and then order it descending based on lymphocyte.spam_matched
    **end if**
    Select top lymphocytes in Innate_Immune_System
**End**

3.3.2.1 Digital antibody

The spam immune system uses string pattern matching to represent this. Pattern recogntion is used in spam detection which means that any given antibody can be

used against more than one infection of spam messages, this is similar to the biological immune system which uses the same antibodies against infection and reinfection.

3.3.2.2 Weights
With each lymphocyte, pieces of information are stored.

- Spam_Matched: the total number or weight of apperance only in spam messages.
- Message_Matched: the total number or weight of apperance of messages by this lymphocyte.

In training, both of these matches are initialized to zero, when the detector matches a message, Message_Matched is incermented by 1, but if that message matched is spam, Spam_Matched will also be incremented by 1.

The two numbers, Spam_Matched and Message_Matched, can be used to give a weighted percentage of the time an antibody detects spam. The field Message _Matched gives an indication of how often this antibody has been used, which helps to determine how important it should be in the final weighting.

In this research the weighted average is applied as follows:

$$\text{Weighted Average} = \frac{\sum_{i=1}^{n} \text{matching\_Lymphocytes(Spam\_Matched)}}{\sum_{i=1}^{n} \text{matching\_Lymphocytes(Message\_Matched)}}$$

where:

matching_Lymphocytes(Spam_Matched): the total number of apperance only in spam messages.
matching_Lymphocytes(Message_Matched): total number of apperance of messages by this lymphocyte.

3.3.3. *Lifecycle of lymphocytes*

The lymphocyte which has the large number of Spam_Matched has the greatest chance to be selected and used against any new message. Also, the lymphocyte which has the lowest number of Spam_Matched has the greatest chance to be culled and replaced with a new acquired lymphocyte.

There are many choices for selecting the update interval (see Algorithm 1) such as the number of messages received, the update interval based on time, the user request, etc. However, in this work the aim of using Genetic Algorithm is to determine the update interval. Also the system gives the selected lymphocytes a chance to fight against infection, however, when any lymphocyte becomes useless, it means that the Spam_Matched value will remain static (no change) and there is a new lymphocyte in (Adaptive_Immune_System) that has Spam_Matched score greater than this lymphocyte. It eventually means that old lymphocytes will be culled and new lymphocytes will be added to the list of Innate_Immune_System.

3.3.4. *Innate immune system*

In the biological immune system, there is an innate immune system that has the capability to defeat infections. Also, in spam immune system there is an innate immune system built from a library and must be able to survive against spam. It is known that there is a great number of lymphocytes in this library and by using this number of lymphocytes which is not small the system will be exhaustive. The best choice to solve this problem is to select the best lymphocytes which have the capability to defeat the largest number of spam. This does not mean that all lymphocytes in the innate immune system are useful and will remain forever. Any new promising lymphocyte created in adaptive immune system will be moved to the innate immune system and will take the place of a useless lymphocyte.

3.3.5. *Adaptive immune system*

In the biological immune system humans can obtain external support in the fight against infection through medicine or other methods. In spam immune system there is an adaptive immune system built from spam messages which contain lymphocytes that do not exist in the innate immune sytem.

3.3.6. *Culling the useless lymphocytes*

As mentioned before, the system suffers from some useless lymphocytes in the innate immune system. So the best solution to keep only useful lymphocytes in the innate immune system is culling. In AIS, it occurs after a parameter defined by the user, and this is not the best solution. In this work, a new creative method is used to perform culling based on a Genetic Algorithm that is used to determine an update interval taking into consideration either the time or the number of messages received by the user. In this work, the Genetic Algorithm is used to determine the culling according to the number of messages received.

3.3.7. *Changing the self*

Interests of users do not remain the same all the time. These interests change according to the new needs created by new circumstances. The user may be interested in something for sometime and after that he may be interested in something else. So the system must be able to change itself. This means that what the system determines as spam must not remain spam all the time. The system must have the capability to allow some spam messages to get into the inbox with time to check if the user becomes intersted in this type of messages or not. If the user becomes interested in this type of messages the system must allow this type of messages to get into the inbox in the future, but if the user continues to have no interest in this type of messages there will be no change on the self. This problem is being solved by adjusting the system to allow some spam message to get into the inbox. Genetic Algorithm is used to determine the rebuild-time which suits any changes in the user's needs to allow these users to read these messages which cover any change on the self.

The success of AIS and genetic optimized AIS spam filtering techniques is determined by classic measures of precision, recall, false positive, and false negative.

***Spam Precision*:** the percentage of messages classified as spam that actually are spam.
***Legitimate Precision*:** the percentage of messages classified as legitimate that are indeed legitimate.
***Spam recall*:** the proportion of the number of correctly-classified spam messages to the number of messages originally categorized as spam.
***Legitimate recall*:** the proportion of correctly-classified legitimate messages to the number of messages originally categorized as legitimate.



## 4. Evaluation and Testing of the Spam Detection Techniques

This part describes the libraries and the parameters used in testing the spam immune system; genetic optimized spam immune system and spam detection using ANN. Experimentation and simulation results are depicted in this section.

### 4.1. *Spam corpus*

The task of selecting a corpus for the evaluation of any learning algorithms of spam detection is difficult. One challenge is that private e-mails are rarely available for public study. In order to test the system, it is necessary to have a public available corpus of e-mails. There are many researchers who use their own personal e-mails or any other available e-mails for training and for the evaluation of their systems. But this will not give broad ranging evidence to prove that the system will work efficiently.

### 4.2. *SpamAssassin public corpus*

The SpamAssassin (SA) corpus is a larger collection made available by spamassassin.org. The SpamAsassin corpus has been used in some research. It contains 4147 legitimate and 1764 (only in 2002) spam messages collected from public or donated by individual users. In this work we use SpamAssassin corpus since it is one of the fewest free pubic corpus and evaluating the system on our private e-mails will not give broad evidence.

### 4.3. *Preparing the corpus*

The corpus was divided depending on the information found in the date field "e-mail header". This information is not necessarily accurate, since it relies on the time of clock of the sender, which may not be correct. All the messages whose data fields were outside 2002 were discarded. The breakdowns of the dataset used in this paper are shown in Table 1.

The SpamAssassin public corpus is divided into seven parts: 20021010 easy ham, 20021010 hard ham, 20030228 easy ham, 20030228 easy ham 2, and 20030228 hard ham contains the non-spam messages. 20030228 spam, 20030228 spam 2, 20050311

Table 1. SpamAssassin public corpus by month.[11,15]

| | Non Spam | | | | Spam | | |
|---|---|---|---|---|---|---|---|
| | Easy Ham | Easy Ham 2 | Hard Ham | Total Non-Spam | Spam | Spam 2 | Total Spam |
| Jan | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| Feb | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Mar | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Apr | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| May | 0 | 0 | 1 | 1 | 0 | 1 | 1 |
| Jun | 0 | 0 | 6 | 6 | 0 | 502 | 502 |
| Jul | 0 | 550 | 152 | 702 | 6 | 566 | 572 |
| Aug | 423 | 843 | 34 | 1300 | 157 | 169 | 326 |
| Sep | 1283 | 0 | 28 | 1311 | 321 | 0 | 321 |
| Oct | 726 | 0 | 9 | 735 | 6 | 0 | 6 |
| Nov | 16 | 6 | 14 | 36 | 0 | 7 | 7 |
| Dec | 52 | 1 | 2 | 55 | 11 | 18 | 29 |



spam 2 and 20021010 spam contains the spam messages. The parts marked 2 indicate more recent additions to the collection. The easy and hard ham indicates which messages have features which make them seem more like spam. Only five parts of the corpus were selected for the study; 20030228 easy ham, 20030228 easy ham 2 and 20030228 hard ham contain the ham messages. Whereas 20030228 spam and 20050311 spam 2 contain the spam messages.

It must be noted that the SpamAssassin public corpus is difficult to be classified and does not reflect a normal ratio of spam to ham. The corpus contain few non-spam messages from January to June, then the number increased during the next months before going back to smaller numbers. It is probably not typical behavior for an individual's mailbox. This pattern is more similar to the way in which the mails were collected by the SpamAssassin public corpus team.

#### 4.3.1. *Training data*

The messages from January to July were chosen to be the training set. It was only in July and June that there were enough non-spam and spam messages in the corpus for sufficient training. Training data contains 710 non-spam messages and 1075 spam messages. The breakdowns are shown in Table 2.

#### 4.3.2. *Testing data*

The messages from August to December were chosen to be the testing set. It was in August and September that there were enough non-spam and spam messages in the corpus for sufficient testing. Testing data contains 3437 non-spam messages and 689 spam messages. The breakdowns are shown in Table 3.

### 4.4. *Spam detection using AIS results*

By testing the system using standard AIS the following results are obtained: (Tables 4–6).

Table 2. Training data by month.[11,15]

| | Non Spam | | | | Spam | | |
|---|---|---|---|---|---|---|---|
| | Easy Ham | Easy Ham 2 | Hard Ham | Total Non-Spam | Spam | Spam 2 | Total Spam |
| Jan | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| Feb | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Mar | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Apr | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| May | 0 | 0 | 1 | 1 | 0 | 1 | 1 |
| Jun | 0 | 0 | 6 | 6 | 0 | 502 | 502 |
| Jul | 0 | 550 | 152 | 702 | 6 | 566 | 572 |
| | | | | 710 | | | 1075 |

Table 3. Testing data by month.[11,15]

| | Non Spam | | | | Spam | | |
|---|---|---|---|---|---|---|---|
| Month | Easy ham | Easy Ham 2 | Hard Ham | Total Non-Spam | Spam | Spam 2 | Total Spam |
| Aug | 423 | 843 | 34 | 1300 | 157 | 169 | 326 |
| Sep | 1283 | 0 | 28 | 1311 | 321 | 0 | 321 |
| Oct | 726 | 0 | 9 | 735 | 6 | 0 | 6 |
| Nov | 16 | 6 | 14 | 36 | 0 | 7 | 7 |
| Dec | 52 | 1 | 2 | 55 | 11 | 18 | 29 |
| | | | | 3437 | | | 689 |

Table 4. Spam precision, recall results (AIS standard).

| No of Lym. | Spam Precision (%) | Spam Recall (%) |
|---|---|---|
| 100 | 57.471 | 65.982 |
| 200 | 81.475 | 66.422 |
| 300 | 82.342 | 67.009 |
| 400 | 82.734 | 67.449 |
| 500 | 84.014 | 69.355 |
| 600 | 90.566 | 70.381 |
| 700 | 93.117 | 71.408 |
| 800 | 93.499 | 71.701 |
| 900 | 92.992 | 71.994 |
| 1000 | 88.470 | 61.877 |

Tables 4 and 5 show spam precision and recall, ham precision and recall, respectively. With a smaller number of lymphocytes the percentage of spam precision and spam recall are low. However, an acceptable value of percentage is achieved when the number of lymphocytes is between 500 and 900.

Based on the results in Table 6, acceptable false positive and false negative rates appear when the number of lymphocytes is between 500 and 900. The best balanced results occur when the number of lymphocytes is 600.

Table 5. Ham precision, recall results (AIS standard).

| No of Lym. | Ham Precision (%) | Ham Recall (%) |
|---|---|---|
| 100 | 93.041 | 90.306 |
| 200 | 93.569 | 97.001 |
| 300 | 93.683 | 97.147 |
| 400 | 93.766 | 97.205 |
| 500 | 94.119 | 97.380 |
| 600 | 94.369 | 98.544 |
| 700 | 94.574 | 98.952 |
| 800 | 94.630 | 99.010 |
| 900 | 94.678 | 98.923 |
| 1000 | 92.857 | 98.399 |

Table 6. False positive, false negative, total error results (AIS standard).

| No of Lym. | False Positive (%) | False Negative | Total Error (%) |
|---|---|---|---|
| 100 | 8.088 | 5.635 | 13.72 |
| 200 | 2.502 | 5.562 | 8.06 |
| 300 | 2.380 | 5.465 | 7.85 |
| 400 | 2.332 | 5.392 | 7.72 |
| 500 | 2.186 | 5.077 | 7.26 |
| 600 | 1.214 | 4.906 | 6.12 |
| 700 | 1.274 | 4.936 | 5.61 |
| 800 | 1.399 | 4.688 | 5.51 |
| 900 | 1.799 | 4.639 | 5.54 |
| 1000 | 2.336 | 6.315 | 7.65 |

Table 7. Spam precision, recall results (genetic optimized).

| No of Lym. | Spam Precision (%) | Spam Recall (%) |
|---|---|---|
| 100 | 97.603 | 65.689 |
| 200 | 93.028 | 74.340 |
| 300 | 92.014 | 77.713 |
| 400 | 78.451 | 81.672 |
| 500 | 96.538 | 69.501 |
| 600 | 91.986 | 77.419 |
| 700 | 93.333 | 73.900 |
| 800 | 93.333 | 73.900 |
| 900 | 93.333 | 73.900 |
| 1000 | 97.863 | 67.155 |

### 4.5. *Genetic optimized AIS spam detection*

#### 4.5.1. *Testing scores*

After testing the system, Tables 7–9 show the following results using a specific threshold:

Getting high values of spam precision means that there are few messages which are incorrectly classified as spam.

Table 8. Ham precision, recall results (genetic optimized).

| No of Lymph. | Ham precision (%) | Ham Recall (%) |
|---|---|---|
| 100 | 93.603 | 99.680 |
| 200 | 95.101 | 98.894 |
| 300 | 95.707 | 98.661 |
| 400 | 96.331 | 95.546 |
| 500 | 94.264 | 99.505 |
| 600 | 95.653 | 98.661 |
| 700 | 95.024 | 98.952 |
| 800 | 95.024 | 98.952 |
| 900 | 95.024 | 98.952 |
| 1000 | 93.861 | 99.709 |

Table 9. False positive, false negative, total error result (genetic optimized).

| No of Lymph. | False Positive (%) | False Negative (%) | Total Error (%) |
|---|---|---|---|
| 100 | 0.267 | 5.684 | 5.95 |
| 200 | 0.923 | 4.251 | 5.17 |
| 300 | 1.117 | 3.692 | 4.81 |
| 400 | 3.716 | 3.036 | 6.75 |
| 500 | 0.413 | 5.052 | 5.47 |
| 600 | 1.117 | 3.741 | 4.86 |
| 700 | 0.874 | 4.324 | 5.20 |
| 800 | 0.874 | 4.324 | 5.20 |
| 900 | 0.874 | 4.324 | 5.20 |
| 1000 | 0.243 | 5.441 | 5.68 |

With fewer numbers of lymphocytes, the error is high for spam recall. This is not true with spam precision rate which remains significantly more constant.

The following tables show false positive and false negative results.

According to Table 9, the false positive rate is low which is more important than the false negative rate. Getting a false negative rate value greater than the false positive value is acceptable for any users since all messages will appear in the user's

Table 10. AIS, genetic optimized AIS summary.

| | AIS Standard | | Genetic Optimized AIS | |
|---|---|---|---|---|
| No of Lymp. | False Positive (%) | False Negative (%) | False Positive (%) | False Negative (%) |
| 200 | 2.502 | 5.562 | 0.923 | 4.251 |
| 300 | 2.380 | 5.465 | 1.117 | 3.692 |
| 400 | 2.332 | 5.392 | 3.716 | 3.036 |
| 500 | 2.186 | 5.077 | 0.413 | 5.052 |
| 600 | 1.214 | 4.906 | 1.117 | 3.741 |
| 700 | 1.274 | 4.936 | 0.874 | 4.324 |
| 800 | 1.399 | 4.688 | 0.874 | 4.324 |
| 900 | 1.799 | 4.639 | 0.874 | 4.324 |
| 1000 | 2.336 | 6.315 | 0.243 | 5.441 |

Table 11. AIS and genetic optimized AIS (average time to test each message).

| | AIS Standard | | | Genetic Optimized AIS | |
|---|---|---|---|---|---|
| No. of Lymp. | Total Number of Messages | Total Time (Hr) | Average Time to Test each Message (Second) | Total Time (Hr) | Average Time to Test each Message (Second) |
| 200 | 4117 | 0.58 | 0.50716541 | 1.02 | 0.891912 |
| 300 | 4117 | 1.09 | 0.9531212 | 1.14 | 0.996842 |
| 400 | 4117 | 1.21 | 1.05805198 | 1.26 | 1.101773 |
| 500 | 4117 | 1.42 | 1.24168084 | 1.51 | 1.320379 |
| 600 | 4117 | 1.58 | 1.38158854 | 2.07 | 1.810056 |
| 700 | 4117 | 2.2 | 1.92373087 | 2.29 | 2.002429 |
| 800 | 4117 | 2.29 | 2.00242895 | 2.35 | 2.054894 |
| 900 | 4117 | 2.4 | 2.0986155 | 2.49 | 2.177314 |
| 1000 | 4117 | 2.55 | 2.22977897 | 2.63 | 2.299733 |

*Note*: All testing is done on PC Pentium Dual Core, 2.5 GHZ, 2GB RAM.



regular mail. Misclassified clean messages are kept in the spam folder, which the user will mostly not open or he/she may delete the messages before seeing them.

## 5. Comparing Standard AIS and Genetic AIS

In this section a comparison of the results are shown (see Table 10) of our experiments on SpamAssassin corpus. You can find that genetic optimized spam detection gives the best results using 600 lymphocytes with 1.117% false positive and 3.741% false negative. Furthermore, spam detection using AIS gives the best results when the number of lymphocytes is 600 with 1.214% false positive and 4.906 false negative.

In the following paragraph, Table 11 shows time taken in testing phase by each algorithm used.

## 6. Discussions

Our Artificial Immune System has successfully captured the biological immune system model and has used it in spam detection. The overall results shown are good enough to be acceptable. Moreover, by using just one single approach (the immune system approach) to achieve such accuracy is really promising. Most of the commercial anti-spam systems use several combined approaches such as the origin based and content filtering methods. A combination of several distinct approaches can achieve higher occurence.

Using GA in genetic optimized spam detection was helpful in determining the culling time and the rebuild time parameters instead of using fixed parameters. This means that GA was the search mechanism to discover different culling time and rebuild time. This is more realistic in real life because users' interests do not change regularly but can be suddenly changed according to any new occurrence.

Concerning the adaptive weight of lymphocytes; it enables the system to replace old lymphocytes with new promising ones. This merit is necessary, useful and helpful

in making the system work better; the results achieved using this approach are promising as it succeeds in modifying the system to work in a similar way to the biological human system. As a result, it can learn new things about spam and therefore, it can develop itself without any intrusion from user/human. However, this advantage cannot guarantee 100% success without running the system for long periods. Testing the system using different numbers of lymphocytes shows that the accuracy can be achieved while using specific numbers of lymphocytes. In addition, it is discovered that when the number of lymphocytes is small, a high percentage of numbers of false positive values appear.



## 7. Conclusions

In this work, the results achieved are generally good for false positive values and acceptable for false negative values (Table 10). Taking into consideration that a balance between a false positive and a false negative must be done, the results summarized in Table 10 demonstrate those acceptable values for both false positives and false negatives. In a few words, genetic optimized spam detection using AIS is good enough to be used as an anti spam effective detection method. The software built is ready to be adopted and tested for commercial purposes. Future work includes using the Artificial Neural Networks (ANN) in enhancing the performance of this hybrid system.

## References

1. I. Androutsopoulos, J. Koutsias, K. Chandrinos, G. Paliouras and C. Spyropoulos, An evaluation of naïve Bayesian antispam filtering, *Proc. Workshop on Machine Learning in the New Information Age, 11th European Conf. Machine Learning (ECML 2000)*, eds. G. Potamias, V. Moustakis, and M. van Someren, (Barcelona, Spain, 9−17, 2000).
2. I. Androutsopoulos, G. Paliouras, V. Karkaletsis, G. Sakkis, C. Spyropoulos and P. Stamatopoulo, Learning to filter spam e-mail: A comparison of a naive Bayesian and a memory based approach, in *Proc. Workshop on Machine Learning and Textual Information Access, 4th European Conf. Principles and Practice of Knowledge Discovery in Databases (PKDD 2000)* (Lyon, France), eds. H. Zaragoza, P. Gallinari and M. Rajman (2000) 1−13.
3. I. Androutsopoulos, J. Koutsias, K. Chandrinos and C. Spyropoulos, An experimental comparison of naïve Bayesian and keyword-based anti-spam filtering with personal e-mail messages, in *Proc. SIGIR, 2000. Ann. ACM Conf. Research and Development in Information Retrieval*, pp. 160−167.
4. E. Batista, A fight to ban cellphone spam, WiredNews, 6 July 2000, available from http://www.wired.com/news/business/0,1283,37376,00.html; "Spam: A New Nuisance for Wireless Users," USA Today, 13 April 2001.
5. J. D. Brutlag and C. Meek, Challenges of the e-mail domain for text classification, *Proc. 17th Int. Conf. Machine Learning* (Stanford University, USA, 2000), pp. 103−110.
6. J. Carpinter and R. Hunt, Tightening the net: A review of current and next generation spam filtering tools, *Comput. Security* **25** (2006) 566−578.
7. X. Carreras and L. Andm'Arquez, Boosting trees for antispam e-mail filtering, *Proc. RANLP-2001, 4th Int. Conf. Recent Advances in Natural Language Processing* (2001).

8. D. Castro and J. Timmis, Artificial immune systems: A novel paradigm to pattern recognition, *Artificial Neural Networks in Pattern Recognition, SOCO-2002*, pp. 67−84.
9. CipherTrust, Inc. Controlling Spam The IronMail® Way, June 2004; http://www.ciphertrust.com/files/forms/landing_template.php?sp.
10. W. W. Cohen, Learning rules that classify e-mail, *Proc. 1996 AAAI Spring Symp. Machine Learning in Information Access*, California.
11. E. Crawford, J. Kay and E. McCreath, IEMS — The Intelligent Email Sorter.
12. D. Dipankar, Advances in artificial intelligence, *IEEE Comput. Intell. Mag.* **1**(4) (2006) 40−49.
13. H. Drucker, D. Wu, and V. N. Vapnik, Support vector machines for spam categorization, *IEEE Trans. Neural Networks* **10**(5) (1999) 1048−1054.
14. K. Gee, Using latent semantic indexing to filter spam, *Proc. 2003 ACM Symp. Applied Computing*, pp. 460−464.
15. C. James, E-mail classification: A hybrid approach combining genetic algorithm with neural networks, PhD Thesis, The University of Sydney, Australia (2000).
16. A. Kolcz and J. Alspector, SVM-based filtering of e-mail spam with content-specific misclassification costs, *Proc. Text DM'01 Workshop on Text Mining 2001 IEEE Int. Conf. Data Mining* (2001).
17. N. Leandro and J. Fernando, Artificial immune systems: Part I — Basic theory and applications, Technical Report TR − DCA 01/99, December, 1999.
18. B. Massey M. Thomure R. Budrevich and S. Long, Learning spam: Simple techniques for freely-available software, *Proc. Usenix Annual Technical Conf., Freenix Track 2003.* http://web.cecs.pdx.edu/∼bart/papers/spam.pdf. Accessed 20 Jun 2008.
19. B. Medlock, A generative, adaptive language model approach to spam filtering, Master Thesis, University of Cambridge MPhil degree in Computer Speech, Text and Internet Technology (July 2003).
20. J. Postel, RFC706: On the Junk Mail Problem, Technical report, Network Working Group, November 1975. http://www.faqs.org/rfcs/rfc706.html.
21. D. Puniškis, R. Laurutis and R. Dirmeikis, An artificial neural nets for spam e-mail recognition, *Electronics and Electrical Engineering* **5**(69).
22. M. Sahami, S. Dumais, D. Heckerman and E. Horvitz, A Bayesian approach to filtering junk e-mail. In Learning for Text Categorization: Papers from the 1998 Workshop. AAAI Technical Report WS-98-05, Madison, WI, 1998; http://www.aaai.org/Library/Workshops/1998/ws98-05-009.php. Accessed 1 Dec. 2008.
23. K. Schneider, A comparison of event models for naive bayes antispam e-mail filtering, *Proc. 11th Conf. European Chapter of the Association for Computational Linguistics (EACL'03)*, 2003; http://scholar.google.com/scholar?hl=ar&lr=&q=Schneider.+K.+2003.+A+comparison+of+event+models+for+naive+bayes+antispam+e-mail+filtering.
24. O. Terri and W. Tony, Increasing the accuracy of a spam-detection immune system, *Proc. Congress on Evolutionary Computation (CEC 2003)*, Vol. 1 (Canberra, Australia, December 2003), 390−396.
25. A. Wiehes, Comparing anti spam methods, Master Thesis, Department of Computer Science and Media Technology, Gjøvik University College, 2005.
26. L. Zhang, J. Zhu and T. Yao, An evaluation of statistical spam filtering techniques, *ACM Trans. Asian Lang. Inform. Process.* **3**(4) (2004) 243−269.

**Raed Abu Zitar** is currently professor of Computer Engineering at New York Institute of Technology, Amman, Jordan. He received his B.S. from the University of Jordan in 1988, MSc. in electrical engineering from North Carolina A&T State University, 1989, and Ph.D. from Wayne State University in 1993.

His research interests include neural networks, machine learning, and pattern recognition.

**Adel Hamdan Mohammed** received his B.S. in computer science, in 1998, from Philadelphia University, Jordan, M.Sc. and Ph.D., in 2005 and 2009, respectively from CIS Arab Academy in banking and financial sciences, Jordan. He is currently an assistant professor at the Computer Science Department, Applied Science University, Jordan.

His areas of expertise are in artificial intelligence, data mining, and pattern recognition.