TOMASZ PRACZYK
Polish Naval Academy

SELECTING GENERATORS FOR CREATING REAL VALUED DETECTORS IN SHIP IMMUNE SYSTEM

ABSTRACT

The task of the ship immune system is to differentiate self objects, i.e. objects that are not dangerous to our ship, from other objects that can be a potential threat. To perform the task the system makes use of a set of detectors. The detectors imitate signatures of non-self objects and they are generated at random. In order for the detectors to be able to effectively perform their task they have to be constructed in appropriate way. Since, random generators are used to form detectors the problem is to select a generator producing the most effective detectors. In order to select an appropriate generator, experiments were carried out. In the experiments, the task of the ship immune system was to differentiate self and non-self ship radio stations. Results of the experiments are presented at the end of the paper.

Keywords:
identification, artificial immune system.

INTRODUCTION

Ship Immune System (SIS) [8] is Artificial Immune System (AIS) [1, 2, 3, 4, 5, 6] whose the main task is to differentiate self ships from non-self ones. The identification of ships is performed based on their signatures, e.g. radio signals generated by ship radio stations. To identify a ship its signature is compared to detectors memorized in SIS. If at least one detector is similar to the signature, the ship is considered to be non-self. Otherwise, it is treated as self.

In SIS, the detectors are generated at random. All detectors which classify self ships as non-self are removed. The remaining detectors are used as 'mature' detectors to identify ships. Generally, in SIS, the detectors can be in the form of real valued, integer valued or binary vectors [8]. In all the cases, the problem is how to generate detectors so as to make SIS possibly the most effective. Since, generating integer valued or binary detectors seems to be simpler than generating real valued ones, in the paper, the problem of generating real valued detectors is raised. The main goal of the paper is to suggest several example detector generators and to select the most effective of them. To this end, experiments were carried out. In the experiments, the task of SIS equipped with detectors generated by means of different generators was to differentiate self radio stations from non-self ones. Results of the experiments are presented at the end of the paper.

The paper is organized as follows: section 2 outlines SIS; section 3 presents detection schemes used in SIS; section 4 proposes several example detector generators; section 5 reports the experiments, and section 6 summarizes the paper.

THE CONCEPT OF SIS

Generally, SIS works as AIS. That is, at first the set of signatures of self ships is created. The signatures from this set are used to create the set of mature detectors of non-self ships. Once the set of self signatures is created the system starts to generate immature detectors. The immature detectors are generated at random. Each immature detector is compared to all self signatures. To survive and to become mature an immature detector has to be different from all self signatures. Otherwise, it is eliminated and replaced with other randomly generated immature detector. The process of generating immature detectors is continued during all the 'life' of SIS. This makes it possible to adapt the system to continuous changes of signatures. Immature detectors which passed the test become mature detectors. The mature detectors participate in the identification of objects. To detect non-self objects the mature detectors use detecting schemes (or matching rules) measuring similarity between the detector and the signature of an unknown object. The process of detecting non-self objects by means of the mature detectors, in detail, is described in the following section. The lifetime of the mature detectors, like their immature counterparts, is not infinite. The mature detectors can also be eliminated.

This can happen in two situations. First, when they are responsible for misclassification of a number of objects in turn. Second, once they are selected for replacement. The replacement of mature detectors with new immature detectors is performed periodically and is necessary in order for the set of mature detectors to include, all the time, up-to-date detectors. The detectors for replacement are selected at random, based on their lifetime, or based on frequency of detections performed by detectors. The simplified model of SIS is presented in figure 1.

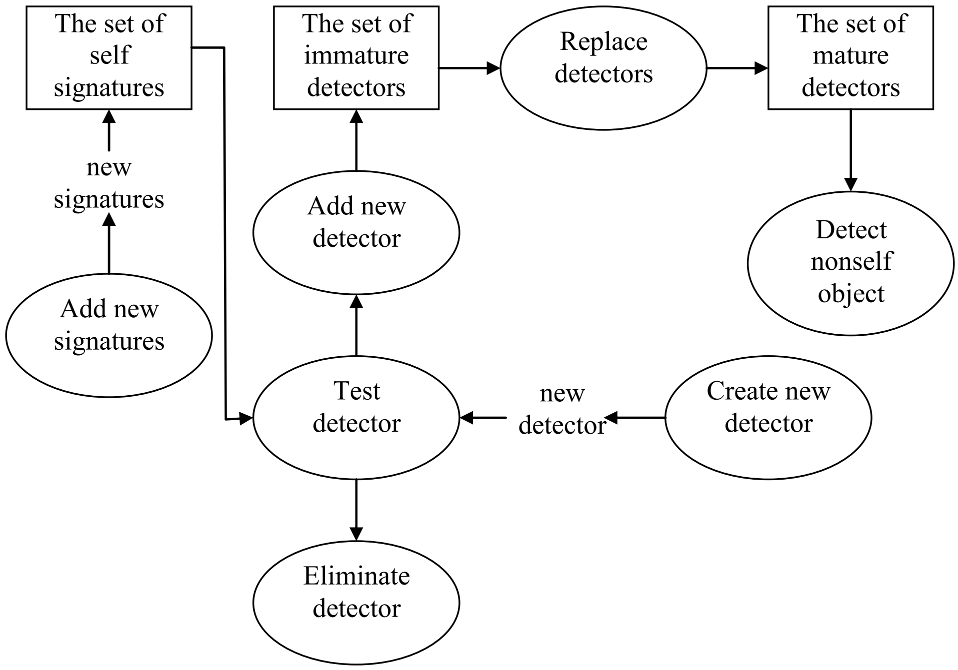

Fig. 1. Model of SIS

DETECTION SCHEMES

To detect non-self objects the mature detectors use detecting schemes. The detection schemes measure similarity between a detector and a signature of an object being identified. Models of AIS created so far use detectors and signatures of objects in the form of binary strings. Usually, to identify non-self objects the models mentioned use the following detecting schemes: Hamming distance, *r*-contiguous-bits rule [3]. However, in the case of ships, we usually deal with signatures in the form of real valued vectors. To detect ships represented in the form of real valued vectors the following detection schemes can be used [7, 8]:

1. Detection scheme No. 1 (Euclidean distance)

$$\mathbf{x}\,\mathsf{M}^{\delta(1)}\mathbf{y} \Leftrightarrow d^{E}(\mathbf{x},\mathbf{y}) \le \delta\,, \tag{1}$$

where:

$\mathbf{x}, \mathbf{y}$ — real valued vectors;

d^E — Euclidean distance;

δ — a parameter;
$\mathbf{x}M^{\delta}\mathbf{y}$ — means that the vectors **x**, **y** match each other.

2. Detection scheme no. 2 (Partial Euclidean distance)

$$\mathbf{x}\,\mathrm{M}_r^{\delta(2)}\mathbf{y} \Leftrightarrow \underset{i}{\exists}\quad \mathbf{x}[i,r]\,\mathrm{M}^{\delta(1)}\mathbf{y}[i,r], \tag{2}$$

where:
x[*i, r*] — a window of size *r* included in the vector **x**; the window begins from the position *i*.

3. Detection scheme No. 3

$$\mathbf{x}\,\mathrm{M}_r^{\delta(3)}\mathbf{y} \Leftrightarrow \underset{i}{\exists}\ \underset{j\in i\ldots i+r}{\forall} \left|\mathbf{x}[j]-\mathbf{y}[j]\right| < \delta, \tag{3}$$

where
x[*i*] — i^{th} element of the vector **x**.

4. Detection scheme No. 4

$$\mathbf{x}\,\mathrm{M}_r^{\delta(4)}\mathbf{y} \Leftrightarrow \underset{i}{\exists}\quad \mathbf{x}\,\mathrm{M}^{\delta(1)}\mathbf{y}[i,r], \tag{4}$$

where:
the vector **x** is of size *r*.

All the schemes specified above were used in the experiments reported further.

GENERATING DETECTORS

In SIS, three elements decide about effectiveness of the system, i.e. type of detectors, parameters of detectors, and the method for generating detectors. The latter element influences the way of locating detectors in the space of ship signatures. Theoretically, the detectors should fill up all the space of ship signatures except areas with concentrations of signatures of self ships. However, in the case of real valued detectors and ship signatures in the form of radio signals, it seems that it is not necessary. The detectors should rather be located only in areas where radio signals can appear. In other places of the space the detectors are unnecessary. Nevertheless, the task of generators of real valued detectors still seems rather difficult. They have to produce detectors which have to surround signatures of self ships all over. Leaving holes in the space where signatures of non-self ships can appear can contribute to great difficulties with right identification of ships.

Generally, to generate real valued detectors many different random generators can be used. Figure 2 presents several example generators which can be applied for that purpose. Each of them generates vectors including elements scaled to the range <0,1>. Example detectors produced by means of the generators from figure 2 are presented in figure 3.

```
genRandomDetector1()
{
for(int i=1;i<=d.length;i++)
      d[i]=rand()/RAND_MAX;
}
```
a)

```
genRandomDetector2(parameter)
{
for(i=1;i<=d.length;i++)
      d[i]=(rand()%(parameter+1))
                  /parameter;
```
b)

```
genRandomDetector3()
{
y=self_signatures.getRandom();
for(i=1;i<=d.length;i++)
      d[i]=1-y[i];
}
```
c)

```
genRandomDetector4()
{
y=self_signatures.getRandom();
for(i=1;i<=d.length;i++)
      d[i]=abs(y[i]-rand()/RAND_MAX);
}
```
d)

```
genRandomDetector5(parameter)
{
y=self_signatures.getRandom();
plus=rand()%2;
noise=rand()/(parameter*RAND_MAX)
for(i=1;i<=d.length;i++)
      if(plus)
            d[i]=y[i]+noise;
      else
            d[i]=y[i]-noise;
if(d[i]>1)
      d[i]=1;
if(d[i]<0)
      d[i]=0;
}
```
e)

Fig. 2. Implementations of generators No. 1–5

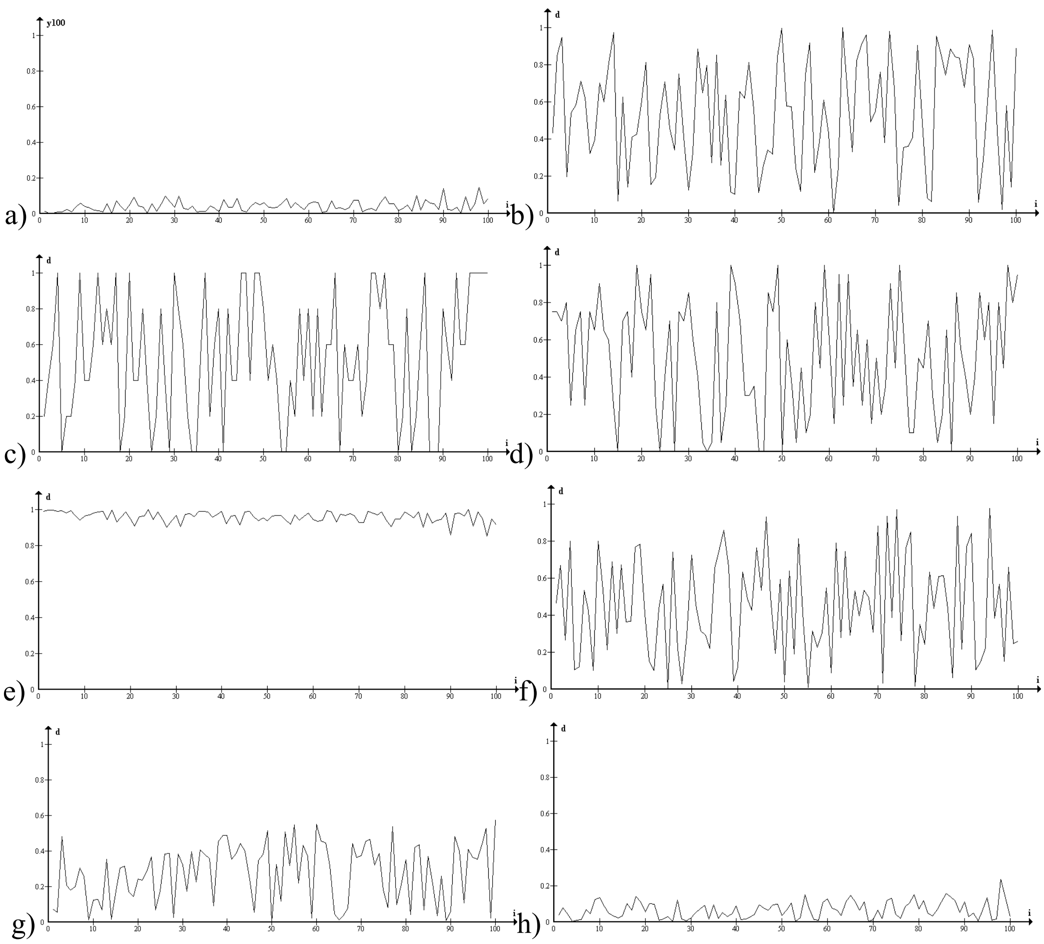

Fig. 3. (a) pattern signature of ship used to generate detectors presented in points (e), (f), (g), and (h); (b) example detector created by generator No. 1 (fig. 2a); (c), (d) example detectors created by generator No. 2 (fig. 2b, parameter = 5 and 20, respectively); (e) example detector created by generator No. 3 (fig. 2c); (f) example detector created by generator No. 4 (fig. 2d); (g), (h) example detectors created by generator No. 5 (fig. 2e, parameter = 2 and 10, respectively)

All the generators specified in figure 2 were tested in the experiments reported further.

EXPERIMENTS

The main goal of the experiments was to select the most effective generator for each detection scheme specified in section 3. In the experiments, ships were represented by radio stations. Accordingly, the task of SIS was to differentiate self radio stations from non-self ones.

Radio signals

In the experiments, ships were represented in the form of encoded radio signals emitted by warship radio stations. Before the signals were used to represent the ships, first, they had been subjected to a feature extraction process. Initially, a discrete spectrum of each signal was fixed. To this end, a discrete Fast Fourier Transform (FFT) was used. Next, a central sample (S_c) of the most informative part of each spectrum was determined. In the following step, vectors including 600 samples to the left and 600 samples to the right from the central sample S_c were created. The vectors were then scaled to the range <0,1>. Since, vectors of size 1200 were still too long to represent ships in SIS (it was very difficult to generate random detectors of size 1200 that would be although slightly similar to any signature of a ship), they were further reduced in size to vectors including 100 samples. To generate signatures of size 100, in the experiments, methods (5)–(8) were used:

$$\mathbf{y}^{100}[i] = \frac{FFT^{1200}[i]}{FFT_{MAX}}, \; i = 1,...,100; \tag{5}$$

$$\mathbf{y}^{100}[i] = \frac{FFT^{1200}[i + 550]}{FFT_{MAX}}, \; i = 1,...,100; \tag{6}$$

$$\mathbf{y}^{100}[i] = \frac{FFT^{1200}[12(i-1) + 1]}{FFT_{MAX}}, \; i = 1,...,100; \tag{7}$$

$$\mathbf{y}^{100}[i] = \frac{FFT^{1200}_{AVERAGE}[12(i-1) + 1,12]}{FFT_{MAX}}, \; i = 1,...,100, \tag{8}$$

where:

$\mathbf{y}^{100}[i]$ — i^{th} element of signature $\mathbf{y}$;

$FFT^{1200}[i]$ — i^{th} sample of radio signal *FFT*;

FFT_{MAX} — maximum value in radio signal *FFT*;

$FFT^{1200}_{AVERAGE}[i,r]$ — average value for a fragment of *FFT* started at sample *i*.

In the experiments, all the methods above were tested in terms of their usefulness in SIS.

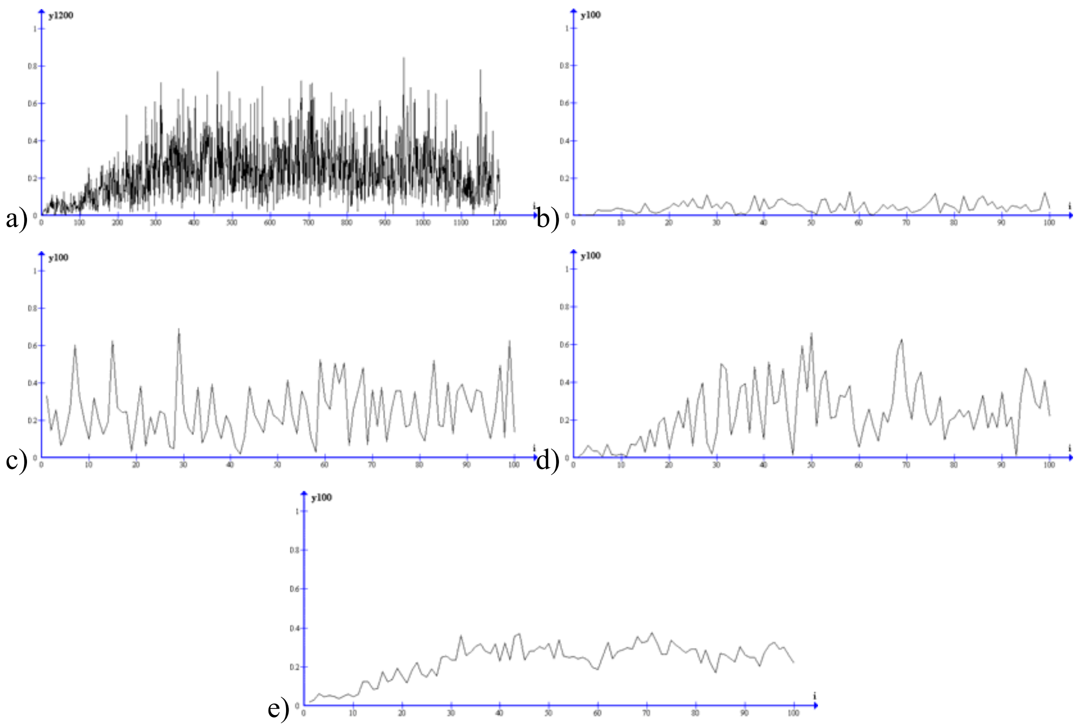

Fig. 4. (a) example signature of size 1200; (b)–(e) signatures of size 100 generated by means of methods (5)–(8) respectively; signatures (b)–(e) generated from signature presented in point (a)

Experimental results

In the experiments, each generator was combined with all the detection schemes and with all the methods used to form signatures of ships. Each combination of the generator, the detection scheme, and the method used to produce ship signatures was tested 30 times. In all the combinations, the following parameters were used: r = 10, number of detectors = 5000. The value of parameter δ was always adjusted to a detection scheme and to a generator. With regard to parameters of detector generators the following values were tested: 5, 10, 20 for generator No. 2, and 2, 5, 10 for generator no. 5.

In the experiments, three sets of radio signals were used. The first set (set No. 1) contained 919 learning signals representing three self warships. It was used to prepare each method. The next set (set No. 2) included 900 signals representing the same three self warships. The set was used to test all the methods. The last set (set No. 3) was composed of 791 signals generated by three warships considered to be non-self. This set was also used to test all the methods specified above.

The results summarizing all the experiments are presented in tables 1–4. Each cell in the tables includes percent of all mistakes (wrong identifications of signals from set No. 2 and 3) made by SIS for a selected combination of the generator, the detection scheme and the method used to form ship signatures. The experiments showed that generator No. 5 with the parameter equal to 2 appeared to be the most suited for schemes (1) and (4). Generator No. 2 with the parameter equal to 5 turned out to be the best for scheme (2). The same generator with the parameter equal to 10 was the best for scheme (3). The experiments also showed that the best solution for all detection schemes is when ships are represented by means of the first hundred samples extracted from original ship signatures of size 1200 (method (5)).

Table 1. Results of experiments for scheme (1) (the best result is bolded)

	method (5)	method (6)	method (7)	method (8)
generator No. 1	47.6%	50.8%	50.2%	50.7%
generator No. 2	47.4% (param=20)	53.8% (param=20)	50.2% (param=20)	52% (param = 20)
generator No. 3	47.1%	46.7%	46.6%	47%
generator No. 4	47.1%	54.4%	59.7%	48.6%
generator No. 5	**30.2% (param=2)**	46.7% (param=2)	46.6% (param=2)	51.4% (param=5)

Table 2. Results of experiments for scheme (2) (the best result is bolded)

	method (5)	method (6)	method (7)	method (8)
generator No. 1	19.5%	65.6%	69.1%	59.7%
generator No. 2	**16.7% (param=5)**	55.4% (param=5)	62.5% (param=5)	54.4%(param=5)
generator No. 3	44%	46.5%	47%	46.8%
generator No. 4	18.9%	52.6%	57.5%	61.5%
generator No. 5	23.1% (param=2)	43% (param=10)	39.9% (param=5)	46.3% (param=5)

Table 3. Results of experiments for scheme (3) (the best result is bolded)

	method (5)	method (6)	method (7)	method (8)
generator No. 1	24.5%	63.8%	66.8%	55.5%
generator No. 2	**19.1% (param=10)**	54.2% (param=5)	69.1% (param=5)	56.7% (param=5)
generator No. 3	45.1%	46.9%	59.6%	47%
generator No. 4	20.1%	57.3%	60.6%	62.6%
generator No. 5	22.7% (param=2)	44.2% (param=5)	52% (param=10)	49.6% (param=5)

Table 4. Results of experiments for scheme (4) (the best result is bolded)

	method (5)	method (6)	method (7)	method (8)
generator No. 1	30.3%	45.7%	53.4%	48.2%
generator No. 2	29% (param=20)	44.6% (param=20)	54% (param=5)	47.7% (param=10)
generator No. 3	41.9%	54.6%	46.8%	46.8%
generator No. 4	28.5%	49%	57.4%	50%
generator No. 5	**19.3% (param=2)**	41.3% (param=2)	54.6% (param=10)	37.5% (param=5)

SUMMARY

The key problem in SIS is a way of generating detectors. In the paper, a few example detector generators were proposed. To test the generators experiments were carried out. In the experiments, the task of SIS was to differentiate self warships from non-self ones. To represent warships, radio signals were used. The experiments showed that two generators should be used in SIS, i.e. generator no. 2 and generator No. 5. The former should be combined with schemes (2), (3), whereas the latter is well suited for schemes (1), (5). In addition to adjusting generators to detection schemes, the experiments also made it possible to determine the most effective representation for ships. In the experiments, it turned out that the best solution for SIS is when ships are represented by means of the first hundred samples extracted from original ship signatures of size 1200.

REFERENCES

[1] Balthrop J., Esponda F., Forrest S., Glickman M., Coverage and Generalization in Artificial Immune System in Proc. Genetic Evolutionary Computation Conf., 2002.

[2] D'haeseleer P., Forrest S., Helman P., An Immunological Approach to Change Detection: Algorithms, Analysis and Implications, Scientific Literature Digital Library, http://citeseer.ist.psu.edu.

[3] Esponda F., Forrest S., Helman P., A Formal Framework for Positive and Negative Detection Schemes, IEEE Transactions on Systems, Man, and Cybernetics — Part B: Cybernetics, 2004, 34(1), pp. 357–373.

[4] Forrest S., Hofmeyr S., Immunology as Information Processing, Design Principles for Immune Systems and Other Distributed Autonomous Systems, eds. Oxford Univ. Press, 2000, pp. 361–387.

[5] Hightower R., Forrest S., Perelson A., The Baldwin Effect in the Immune System: Learning by Somatic Hypermutation, R. K. Belew and M. Mitchell, editors, Individual Plasticity in Evolving Populations: Models and Algorithms, Addison — Wesley, 1996, pp. 159–167.

[6] Hofmeyr S., Forrest S., Somayaji A., Intrusion Detection using Sequences of System Calls, Scientific Literature Digital Library, http://citeseer.ist.psu.edu.

[7] Praczyk T., Adaptation of r-contiguous-bits scheme borrowed from immune systems to characteristic points of radar image identification, Theoretical and Applied Informatics, 2007, Vol. 19, pp. 37–56.

[8] Praczyk T., The concept of the ship immune system, Annual of Navigation, 2009, No. 15, pp. 101–107.

Received April 2009
Reviewed January 2010