Computing and Informatics, Vol. 29, 2010, 975–987

USING REAL VALUED DETECTORS IN SHIP IMMUNE SYSTEM

Tomasz Praczyk

Polish Naval Academy Śmidowicza 69
Gdynia, Poland
e-mail: T.Praczyk@amw.gdynia.pl

Manuscript received 2 April 2009; revised 11 September 2009
Communicated by Vladimír Kvasička

Abstract. The paper addresses the problem of real valued detectors in ship immune system. The task of the system mentioned 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 this end, mechanisms adapted from artificial immune systems are used. Since in the traditional model of artificial immune system binary strings are used to represent detectors and objects, in this paper modifications to this model are proposed. The modifications mentioned use real valued vectors instead of binary ones. To test the ship immune system equipped with real valued detectors, experiments were carried out. In the experiments, the task of the system was to differentiate self ship radio stations from non-self ones. Results of the experiments are presented at the end of the paper.

Keywords: Identification, artificial immune system

Mathematics Subject Classification 2000: 68T05, 68T10, 92D25

1 INTRODUCTION

The "friend or foe" identification is a very important problem during a war. To date, mistakes occur on a battlefield, when self units are destroyed by other self units. At sea, the problem is to identify a submerged submarine. Wrong identification of the ship can have two serious consequences. First, the ship can be self however regarded as non-self. As a result, it can be destroyed by self submarine killers. Second, the ship can be non-self but considered to be a self. In this case, it can be unpunished in fighting against our fleet.

Existing military "friend or foe" systems work in the active way. Thus, to identify a ship exchange of information between ships is necessary. The paper suggests another solution. The solution proposed is based on the idea adapted from artificial immune systems (AISs) [1, 3, 4, 6, 7, 9] and for that reason it has been called Ship Immune System (SIS) [11]. In SIS, a ship is identified in the passive way. Moreover, the identification is performed based on a signature of a ship which can be treated as fingerprints very difficult to counterfeit. The next advantage of SIS is its construction which is based exclusively on signatures of self ships. To build the system, the information about alien ships is not necessary. The system should detect all that differs from signatures of self ships memorized in the system. The next beneficial feature of SIS is its adaptability to changing conditions. This means that any change in the own fleet does not entail the necessity of building a new system. The set of detectors imitating alien ships should always adjust to changing signatures of self ships.

Models of AISs created so far use signatures of objects in the form of binary strings. However, ships are usually represented in the form of real valued vectors (e.g. radio signals generated by ships). For this reason, to detect non-self ships, other than classic detection schemes have to be used. In this paper, several real valued schemes are proposed. The schemes are modifications to different binary schemes. In order to test the schemes, experiments were carried out. In the experiments, the task of SIS, equipped with real valued detectors, was to detect alien ship radio stations. To compare SIS with other methods, in the experiments Probabilistic Neural Network (PNN) [12] and k nearest neighbor method (kNN) were also used.

The paper is organized as follows: Section 2 outlines SIS; Section 3 reports the experiments, and Section 4 summarizes the paper.

2 THE CONCEPT OF SIS

SIS works like AIS. At first the set of signatures of self ships is created. The signatures from this set are used to create the so-called mature detectors imitating antibodies from the natural immune system. Once the set of self signatures is created the system starts to generate immature detectors which do not take part in detection process. 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 by another randomly generated immature detector. The process of generating the immature detectors is continued during the whole "life" of SIS. This makes 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 a detector and a signature of an unknown object. The process of detecting non-self objects by means of the mature detectors is described in detail in the following section. The lifetime of the mature detectors 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 the mature detectors with new immature detectors is performed periodically and is necessary in order for the set of mature detectors to include up-to-date detectors all the time. The detectors for the replacement are selected at random, based on their lifetime, or based on the frequency of detections performed by detectors. The simplified model of SIS is presented in Figure 1.

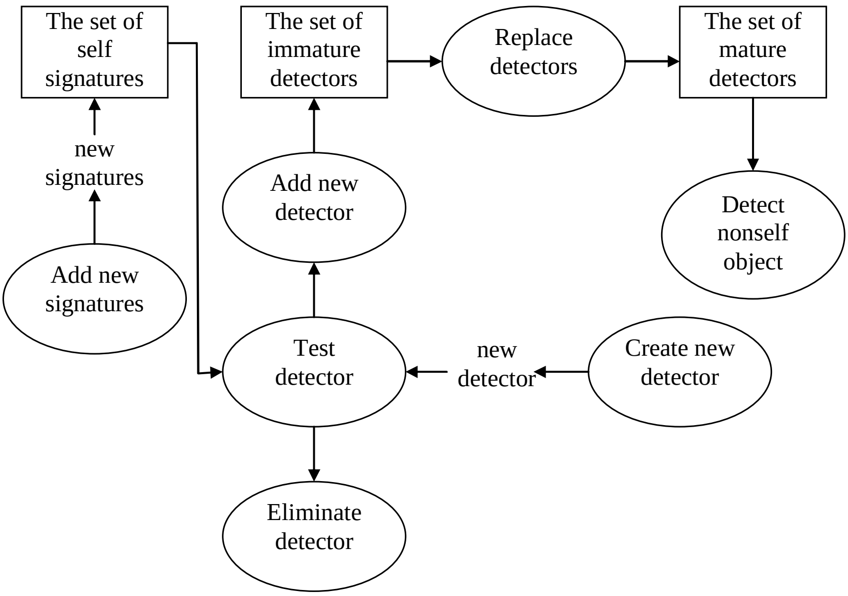

Fig. 1. Model of SIS

2.1 Detection Schemes Used in SIS

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 [4].

a)	b)
11***11***10***01***0001	11111***00100***01
10***11***01***01***1110	10110***00100***11

Fig. 2. a) Using Hamming distance with $r = 5$; strings agree on five bits b) using *rcb* rule with $r = 5$; strings agree on five contiguous bits

In both detecting schemes mentioned above, the result of matching between two strings depends on the value of r. If r equals the length of both strings then any antibody string can only recognize a single antigen string. In turn, if $r = 0$ then each detector string matches each antigen string. Generally, higher values of r make a detecting scheme more specific. In turn, lower values make it more general.

The value of r has influence on discrimination errors, which AIS can make during work. If a self string is identified as foreign then we deal with the so-called false positive. In turn, a false negative takes place when a non-self string is classified as normal. Both errors are dangerous. In the first case, the system attacks oneself, while in the second case it does nothing in order to defend oneself against outside threat.

Above, the detection schemes are discussed which can be used to compare binary strings. However, in the case of ships we deal with signatures in the form of real valued vectors. There are two solutions to this problem. First, we can try to build the system based on the real valued vectors. Second, the real valued vectors representing ships can be reduced to binary strings or integer vectors and to compare them detecting schemes presented above can be used. In the paper, the former solution is considered.

In SIS, and generally in all types of AIS, the very important issue is to completely surround areas containing self objects by detectors. Using the real valued vectors as detectors and signatures of ships can make it difficult. To surround all self signatures, the number of detectors has to be very large. This, in turn, can make the system very slow and, in consequence, useless. The panacea to this problem is to maximally compress the vectors representing antigens and antibodies, and to perform calculations in many locations (many copies of SIS, each copy with its own database including detectors). When the distribution of calculations is impossible the system can also be speeded up through appropriate organization of the set of detectors. The whole set of detectors can be divided into parts. Each part should be represented by a single average detector. To determine average detectors, Kohonen neural network can be used. In this case, the detection process starts from comparing an unknown object with all the average detectors. In the following phase of the detection process, only detectors included in the zone of responsibility of the average detector being the closest to the object being identified are used. The remaining detectors do not take part in the detection process. If some detector from the set of selected detectors is similar to the object being identified the object is treated as non-self.

In the case of real valued vectors, the following detection schemes can be used [10, 11]:

1. Detection scheme No. 1 (Euclidean distance):

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

   where

   - $\mathbf{x}$, $\mathbf{y}$ are real valued vectors;
   - d^E is Euclidean distance;
   - $\mathbf{x} M^{\delta} \mathbf{y}$ means that the vectors $\mathbf{x}$, $\mathbf{y}$ match each other.

2. Detection scheme No. 2 (Partial Euclidean distance):

$$\mathbf{x} M_r^{\delta(2)} \mathbf{y} \Leftrightarrow \exists_i \mathbf{x}[i, r] M^{\delta(1)} \mathbf{y}[i, r] \tag{2}$$

   where

   - $\mathbf{x}[i, r]$ is a window of size r included in the vector $\mathbf{x}$;
   - the window begins from the position i.

3. Detection scheme No. 3:

$$\mathbf{x} M_r^{\delta(3)} \mathbf{y} \Leftrightarrow \exists_i \forall_{j \in i \ldots i+r} |\mathbf{x}[j] - \mathbf{y}[j]| < \delta \tag{3}$$

   where

   - $\mathbf{x}[i]$ is the i^{th} element of the vector $\mathbf{x}$.

4. Detection scheme No. 4:

$$\mathbf{x} M_r^{\delta(4)} \mathbf{y} \Leftrightarrow \exists_i \mathbf{x} M^{\delta(1)} \mathbf{y}[i, r] \tag{4}$$

   where

   - the vector $\mathbf{x}$ is of size r.

5. Detection scheme No. 5:

$$\mathbf{d} M_r^{\delta(5)} \mathbf{y} \Leftrightarrow d_{\#}^E(\mathbf{d}, \mathbf{y}) \leq \delta \tag{5}$$

   where

   - $\mathbf{d}$ is a detector including real valued elements and the so-called "don't care symbols" denoted as "#" (don't care symbol is interpreted as any real value), and

- $d^E_{\#}$ is defined below:

$$d^E_{\#}(\mathbf{d},\mathbf{y}) = \sqrt{\sum_{i,\mathbf{d}[i]\neq`\#'} (\mathbf{d}[i]-\mathbf{y}[i])^2}. \tag{6}$$

In schemes (1)–(4), the detectors are represented in the form of real valued vectors. In scheme (5), the detectors have similar construction to classifiers from Learning Classifier Systems [2, 5, 8], i.e. they include real valued elements and "don't care" symbols. All types of the detectors were tested in the experiments reported in the further part of the paper.

```
genRandomDetector1(parametr,l)
    {
    signatureOfSelfObject=get random signature of size l;
    for i=1 to l;
        {
        noise=generate random value from <0,1>;
        noise=noise/parameter;
        detector[i]=signatureOfSelfObject[i]±noise;
        }
    }
```

Fig. 3. Generator No. 1 used to produce detectors for schemes (1), (4)

2.2 Generating Detectors in SIS

The key problem in SIS is the way of generating detectors. The detectors are created at random and to generate them different random generators can be used. To determine the best generators for each detection scheme specified in section 2.1 preliminary experiments were carried out. Pseudocode of the best generators and example detectors generated by them are presented in Figures 3, 4, 5 and 6. All the generators create detectors of length l.

```
genRandomDetector2(parametr,l)
    {
    for i=1 to l;
        {
        rand=generate random integer from 0 to parameter;
        detector[i]=rand/parameter;
        }
    }
```

Fig. 4. Generator No. 2 used to produce detectors for schemes (2), (3)

3 EXPERIMENTS

The main goal of the experiments was to test whether SIS equipped with real valued detectors can be an effective tool for detecting alien ships. In the experiments,

```
genRandomDetector3(parametr,l)
    {
    for i=1 to l;
        {
        dont_care=generate random integer from 0 to parameter;
        if (dont_care is equal to 0)
            detector[i]=-1 //'#';
        else
            detector[i]=generate random value from <0,1>;
        }
    }
```

Fig. 5. Generator No. 3 used to produce detectors for scheme (5)

SIS was compared to PNN and kNN methods. To represent ships, radio signals generated by ship radio stations were used. Originally recorded signals were converted into vectors including 1 200 and 100 samples. While vectors of size 100 were used in the experiments with all the compared methods, i.e. with SIS, PNN and kNN, vectors of size 1 200 were only applied to build different variants of PNN

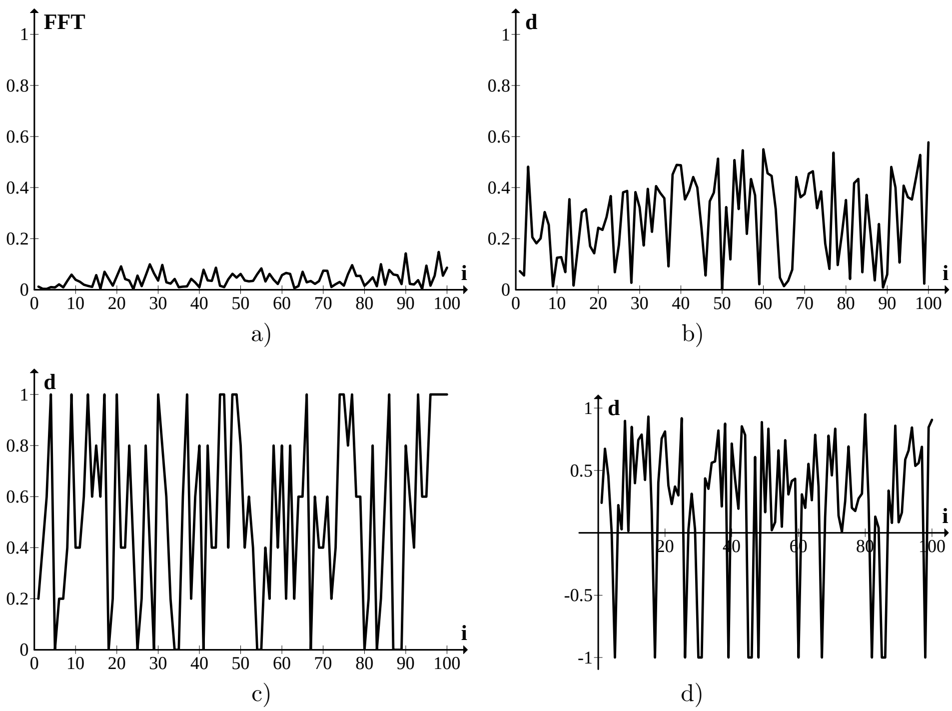

Fig. 6. a) pattern signature of ship used to generate the detector presented in point b), b) example detector created by generator No. 1 (parameter = 2); c) example detector created by generator No. 2 (parameter = 5); d) example detector created by generator No. 3 (parameter = 5)

and kNN. In the case of SIS, signatures of ships of size 1 200 appeared to be definitely too long. It was very difficult to generate random detectors of size 1 200 that would be although slightly similar to any signature of a ship of the same size.

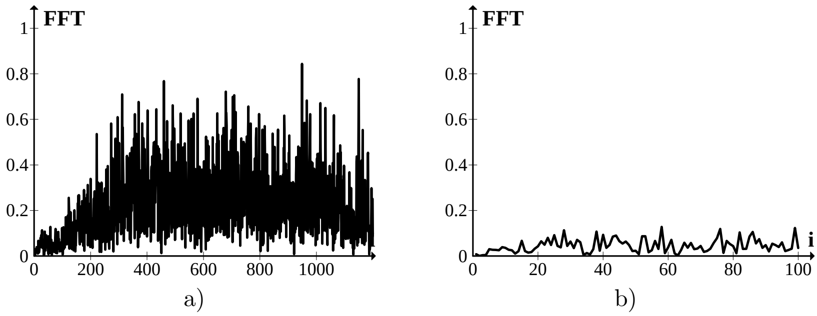

Fig. 7. a) example signature of ship of size 1 200; b) example signature of ship of size 100

In the experiments, the following methods were tested:

1. PNN_1200 – PNN built with original ship signatures of size 1 200;
2. PNN_100 – PNN built with signatures of size 100;
3. 1NN_1200_(1) – kNN with $k = 1$, with detectors of size 1 200, and with detection scheme (1);
4. 2NN_1200_(1) – kNN with $k = 2$, with detectors of size 1 200, and with detection scheme (1);
5. 4NN_1200_(1) – kNN with $k = 4$, with detectors of size 1 200, and with detection scheme (1);
6. 1NN_100_(1) – kNN with $k = 1$, with detectors of size 100, and with detection scheme (1);
7. 2NN_100_(1) – kNN with $k = 2$, with detectors of size 100, and with detection scheme (1);
8. 4NN_100_(1) – kNN with $k = 4$, with detectors of size 100, and with detection scheme (1);
9. 1NN_100_(2) – kNN with $k = 1$, with detectors of size 100, and with detection scheme (2);
10. 1NN_100_(3) – kNN with $k = 1$, with detectors of size 100, and with detection scheme (3);
11. SIS_(1) – SIS with detection scheme (1);
12. SIS_(2) – SIS with detection scheme (2);
13. SIS_(3) – SIS with detection scheme (3);

14. SIS_(4) – SIS with detection scheme (4);
15. SIS_(5) – SIS with detection scheme (5).

All the methods were tested for different values of their parameters. As for PNN and its parameter σ determining a shape of radial functions assigned to each neuron, it appeared that the best value for this parameter is 6.7 for PNN_1200 and 0.016 for PNN_100. In kNN and SIS, parameters δ and optionally r had to be fixed. With regard to δ its value was always experimentally adjusted both to a method and to the value of the parameter r. As for r, the following values were tested: 20, 40, 60, 80 for 1NN_100_(2) and 1NN_100_(3), and 10, 20, 40 for all variants of SIS. In addition to δ and r, the influence on performance of SIS has also the number of detectors. In the experiments, each variant of SIS was tested for 2 000, 5 000 or 10 000 detectors.

To test the methods, 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 next three warships considered to be non-self. This set was also used to test all the methods specified above.

In the experiments, each variant of PNN was tested many times. Each single run of the network was connected with a different value of the parameter σ. To find optimal values of parameters for each variant of kNN, many runs were necessary as well. In a single run one combination of parameters δ and r was tested. In the case of SIS, the first activity was to find the best value of δ for each combination of the number of detectors and the value of r. One run was performed for each tested value of δ. In the next step, all the best configurations of SIS with fixed value of δ were run thirty times (thirty runs were performed for each variant of SIS with a different number of detectors and different r).

3.1 Experimental Results

The results summarizing all the experiments are presented in Table 1. The table includes only the best results (in the case of SIS the results are averaged) for each method tested in the experiments. Generally, the experiments showed that using real valued detectors can be effective method to differentiate self and non-self objects represented as real valued vectors. SIS_(2) turned out to be the best solution out of all the methods tested in the experiments. The mentioned variant of SIS made the least mistakes when identifying ships (all in all 30.2 false positives – 3.3 %, 252.4 false negatives – 31.9 %, and 282.6 of all mistakes – 16.7 %, on average). The schemes (3), (4), and (5) appeared to be somewhat less effective. Nevertheless, they were still better than most methods compared to SIS. The worst result of all was achieved by scheme (1).

With regard to the influence of the number of detectors on performance of SIS, it appeared that, except the detection scheme (4), augmenting the set of detectors over

	% of false positives	% of false negatives	% of all mistakes	parameters
SIS_(2)	3.3 %	31.9 %	16.7 %	$\delta = 0.6$, $r = 10$
1NN_100_(1)	3.3 %	35.2 %	18.2 %	$\delta = 0.65$
2NN_100_(1)	3.4 %	35.1 %	18.2 %	$\delta = 0.65$
4NN_100_(1)	3.6 %	35 %	18.3 %	$\delta = 0.65$
SIS_(5)	2.7 %	36.1 %	18.3 %	$\delta = 6.8$
SIS_(3)	2.6 %	37.9 %	19.1 %	$\delta = 0.3$, $r = 10$
SIS_(4)	4.6 %	36 %	19.3 %	$\delta = 0.4$, $r = 10$
PNN_1200	2.5 %	39.4 %	19.8 %	$\sigma = 6.7$
1NN_100_(2)	7.1 %	34.7 %	20 %	$\delta = 0.33$, $r = 80$
2NN_1200_(1)	2.7 %	40.8 %	20.5 %	$\delta = 4\,800$
4NN_1200_(1)	2.2 %	42 %	20.8 %	$\delta = 4\,900$
1NN_100_(3)	6.7 %	37.2 %	21 %	$\delta = 0.1$, $r = 80$
1NN_1200_(1)	3.2 %	41.4 %	21.1 %	$\delta = 4\,730$
PNN_100	39.4 %	6.9 %	24.2 %	$\sigma = 0.016$
SIS_(1)	8 %	52.4 %	28.8 %	$\delta = 2.4$

Table 1. Results of experiments (methods are ordered according to the fourth column, i.e. from the best to the worst method; % of false positives – self from set No. 2 considered to be non-self; % of false negatives – non-self from set No. 3 considered to be self; % of all mistakes – wrong identifications of signals from set No. 2 and 3)

2 000 does not entail considerable increase in SIS performance (Table 2). Usually, using 2 000 or 5 000 detectors is equally effective as or even more effective than using 10 000 detectors.

number of detectors	scheme (1)	scheme (2)	scheme (3)	scheme (4)	scheme (5)
2 000	29.69 %	17.21 %	19.46 %	23.65 %	18.39 %
5 000	29.63 %	17.09 %	19.40 %	19.34 %	19.22 %
10 000	28.80 %	17.45 %	19.16 %	19.46 %	19.40 %

Table 2. Percentage of all mistakes for different number of detectors

The experiments showed that lower values of r yield better results than larger values. Of course, r cannot be too low, because it can lead to great difficulties with correct identification of ships. In the experiments, three values of r were tested, i.e. 10, 20, and 40. As presented in Table 3, the best results, regardless of the scheme, were achieved for r equal to 10.

In the experiments, it turned out that in order for SIS to be possibly the most effective the value of δ has to be near the maximum value (Tables 4 and 5). The maximum value for δ is the value for which it is not possible or at least it is very difficult to generate any detector. Most of the detectors generated at random are similar to self objects and consequently they cannot become mature and be used to

r	scheme (2)	scheme (3)	scheme (4)
10	16.68 %	19.04 %	19.04 %
20	21.11 %	26.26 %	21.35 %
40	41.81 %	45.48 %	22.53 %

Table 3. Percentage of all mistakes for different r

identify objects.

δ	all mistakes	false positives	false negatives
0.2	46.66 %	0 %	99.75 %
0.4	30.51 %	3.67 %	61.06 %
0.6	17.09 %	4.11 %	31.86 %
0.65	18.33 %	6.67 %	31.61 %

Table 4. The influence of δ on performance of SIS_(2) with $r = 10$ and with 5 000 detectors

δ	all mistakes	false positives	false negatives
0.1	30.46 %	2.89 %	61.82 %
0.2	22.59 %	2.11 %	45.89 %
0.3	19.75 %	2.22 %	39.70 %

Table 5. The influence of δ on performance of SIS_(3) with $r = 10$ and with 5 000 detectors

The last observation from the experiments involves the ship identification itself. It appeared that efficient identification of ships exclusively based on radio signals is very difficult. To identify ships with greater certainty, using radio signals as the only representation for ships seems to be insufficient. To make the identification more reliable, using many different representations (e.g. radar signal, sound generated by ship devices, magnetic field generated by ships) is rather inevitable.

4 SUMMARY

In the paper, SIS equipped with different types of real valued detectors is presented. To test SIS, 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. In addition to SIS, for the purposes of comparison different variants of PNN and kNN were also tested in the experiments. The experiments showed that SIS with the real valued detectors is quite effective solution to the self/non-self detection problem. In the experiments, the variant of SIS with the detection scheme (2) appeared to be the most effective detection method. However, results achieved by the mentioned variant of SIS seem to be further unsatisfactory, being particularly unacceptable when identifying non-self ships – 31.9 % of misclassifications. This result means that almost every third non-self ship would be misclassified. For the military system, working in real conditions, it is definitely too much.

Since the results achieved by SIS seem still unsatisfactory, further experiments are planned. One of the elements which can improve effectiveness of SIS is so-called "vaccination". In SIS, vaccination means introducing non-self objects, i.e. representations of non-self ships, to the system. The objects mentioned are used to generate a part of detectors. Since the new detectors are generated in areas where a great chance to find non-self object occurs, they can reduce the number of false negatives made by SIS.

Performance of SIS can be further enhanced by using detectors in the form of integer or binary strings. Such form of detectors is usually used in traditional models of AIS and for that reason it should also be tested in SIS.

REFERENCES

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

[2] Butz, M. V.: Rule-Based Evolutionary Online Learning Systems: Learning Bounds, Classification, and Prediction. University of Illinois, IlliGAL Report No. 2004034, 2004.

[3] 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`.

[4] 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, Vol. 34, 2004, No. 1, 2004, pp. 357–373.

[5] Goldberg, D. E.: Genetic Algorithms in Search, Optimization and Machine Learning. Addison Wesley, Reading, Massachusetts 1989.

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

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

[8] Holland, J. H.: Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor, Michigan 1975.

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

[10] PRACZYK, T.: Adaptation of r-Contiguous-Bits Scheme Borrowed from Immune Systems to Characteristic Points of Radar Image Identification. Theoretical and Applied Informatics, Vol. 19, 2007, pp. 37–56.

[11] PRACZYK, T.: The Concept of the Ship Immune System. Annual of Navigation (to appear).

[12] SPECHT, D. F.: Probabilistic Neural Networks. Neural Networks, Vol. 3, 1990, No. 1, pp. 109–118,

Tomasz PRACZYK is a senior lecturer at the Institute of Naval Weapon of Polish Naval Academy in Gdynia. He received his M. Sc. degree in computer science in 1996. In 2001, he received his Ph. D. degree; with thesis focused on using artificial neural networks to identify ships. His research interest is in neuro-evolution, artificial immune systems, and reinforcement learning.