RE-EDITING AND CENSORING OF DETECTORS IN NEGATIVE SELECTION ALGORITHM

X. Z. Gao, S. J. Ovaska, and X. Wang

Department of Electrical Engineering
Helsinki University of Technology
Otakaari 5 A, FI-02150 Espoo, Finland
Tel: +358 9 451 2434, Fax: +358 9 451 2432
http://powerelectronics.tkk.fi/
E-mail: gao@cc.hut.fi, seppo.ovaska@tkk.fi, xiaolei@cc.hut.fi

ABSTRACT

The Negative Selection Algorithm (NSA) is a kind of anomaly detection method inspired by the biological self/nonself discrimination principles. In this paper, we propose two new schemes for the detectors re-editing and censoring in the NSA. The detectors that fail to pass the negative selection phase are re-edited and updated to become qualified using the Differential Evolution (DE) method. In the detectors censoring, the qualification of all the detectors is evaluated, and only those appropriate ones are retained. Prior knowledge of the anomalous signals is utilized to discriminate the detectors so that their anomaly detection performances can be improved. The effectiveness of our detectors re-editing and censoring approaches is examined with both artificial signals and a practical bearings fault detection problem.

1. INTRODUCTION

Natural immune systems are complex and enormous self-defense systems with the remarkable capabilities of learning, memory, and adaptation [1]. Artificial Immune Systems (AIS), inspired by the natural immune systems, are an emerging kind of soft computing methods [2]. With the features of pattern recognition, anomaly detection, data analysis, and machine learning, the AIS have recently gained considerable research interest from different communities [3]. As an important constituent of the AIS, Negative Selection Algorithm (NSA) is based on the principles of maturation of T cells and self/nonself discrimination in the biological immune systems. It was firstly developed by Forrest *et al.* in 1994 for the real-time detection of computer viruses [4]. During the past decade, the NSA has been widely applied in numerous interesting engineering areas, e.g., networks security [5] and milling tool breakage detection [6]. The NSA detectors are first generated in a random manner, and undergo the so-called 'negative selection' process thereafter. Only the detectors that do not match the *self* are selected for the anomaly detection, and those unqualified ones will be eliminated. However, practical generation and implementation/manufacture of the detectors can be costly. Therefore, how to re-use the unqualified detectors that are already generated is an important issue, particularly in the expense-sensitive cases. Another drawback of the original NSA is that it is difficult if not impossible to explicitly embed the prior information of the anomaly to be detected into the detectors selection phase. In this study, we first present a Differential Evolution (DE)-based detectors re-editing scheme. A novel method of utilizing the characteristics of the anomalous signals for censoring the NSA detectors is also proposed and explored.

The remainder of this paper is organized as follows. We introduce the essential principles of the NSA in Section 2. The detectors re-editing and censoring approaches are proposed and discussed in Sections 3 and 4, respectively. We explain in details how to employ the DE method to re-edit the unqualified NSA detectors as well as utilize the domain knowledge to censor the coarse detectors. Simulations of three numerical examples of artificial signals and bearings fault detection are made in Section 5 for examining our detectors re-editing and censoring scheme. Finally, in Section 6, we conclude this paper with some remarks and conclusions.

2. Principle of Negative Selection Algorithm

It is well known that the natural immune system is an efficient self-defense system that can protect the human body from being affected by foreign antigens or pathogens [1]. One of its most important functions is pattern recognition and classification. In other words, the biological immune system is capable of distinguishing the self, i.e., normal cells, from the nonself, such as bacteria, viruses, and cancer cells. This capability is mainly achieved by two different types of lymphocytes: B cells and T cells. Both the B cells and T cells are produced in the bone marrow. However, for the T cells, they must pass through a *negative* selection procedure in the thymus thereafter. Only those that do not match the self proteins of the body will be released out to circulate. The remaining others are eventually destroyed there, which can actually prevent our immune system from mistakenly attacking the body's own proteins.

The NSA is inspired by the aforementioned T cell maturation mechanism of the biological immune system, as shown in Fig. 1. This approach can be conceptually described as follows. Defining the self, we first collect a data set containing all the representative self samples. Next, the candidate detectors are *randomly* generated, and compared with the self set. Note that like the above negative selection of the T cells, only those detectors that do not match any element of the self sample set are retained. Let $\left[x_1, x_2, \cdots, x_L\right]$ and $\left[w_1, w_2, \cdots, w_L\right]$ denote a self sample and a candidate detector, respectively, where L is their common order. The matching degree d between $\left[x_1, x_2, \cdots, x_L\right]$ and $\left[w_1, w_2, \cdots, w_L\right]$ can be calculated based on the Euclidean distance:

$$d = \sqrt{\sum_{i=1}^{L} \left(x_i - w_i\right)^2} . \tag{1}$$

d is then compared with a preset threshold λ, and the detector matching error E is obtained:

$$E = d - \lambda . \tag{2}$$

If $E > 0$, detector $\left[w_1, w_2, \cdots, w_L\right]$ fails to match self sample $\left[x_1, x_2, \cdots, x_L\right]$. If $\left[w_1, w_2, \cdots, w_L\right]$ does not match all the self samples, it will be included in the detector set. On the other hand, if $E \le 0$, we consider that $\left[w_1, w_2, \cdots, w_L\right]$ matches $\left[x_1, x_2, \cdots, x_L\right]$, and it is therefore rejected. After a certain number of qualified detectors have been generated by such a negative selection procedure, they are used to detect the nonself or anomaly in the incoming samples. That is, when a new sample $\left[x_1', x_2', \cdots, x_L'\right]$ matches $\left[w_1, w_2, \cdots, w_L\right]$, the existing anomaly is detected. Unfortunately, conventional NSA has the shortcoming of inefficiency in detectors generation [7]. A few modified versions of the NSA have been investigated during the recent years [8]-[11]. However, most of these algorithms just neglect the re-use of the unqualified detectors, and they cannot fully utilize the prior domain information of the anomalous signals. We propose the following detectors re-editing and censoring schemes in the NSA to achieve improved anomaly detection performances.

3. Detectors Re-editing in Negative Selection Algorithm

A. Differential Evolution Method

The Differential Evolution (DE) method is a robust population-based optimization technique firstly proposed by Storn and Price [12]. The principle of the DE is similar to that of other evolutionary programming methods, such as the Genetic Algorithms (GA) [13]. However, the unique idea of the DE is that it generates new chromosomes by adding the weighted difference between two chromosomes to the third one. If the fitness of the resulting chromosome is better than the original chromosome, this newly generated chromosome replaces the one with which it is compared. The simplest DE can be explained as follows. Suppose there are three chromosomes, $r_1(k)$, $r_2(k)$, and $r_3(k)$, in the current population, as shown in Fig. 2. A trial update of $r_3(k)$, $r_3'(k+1)$, is given:

$$r_3'(k+1) = r_3(k) + \lambda\left[r_1(k) - r_2(k)\right], \tag{3}$$

where λ is a pre-determined weight. In order to further increase the diversity of the chromosomes, a 'crossover' operator is employed to generate $r_3''(k+1)$ by randomly combining those parameters of $r_3(k)$ and $r_3'(k)$ together. If $r_3''(k+1)$ yields a higher fitness than $r_3(k)$, we get:

$$r_3(k+1) = r_3''(k+1) . \tag{4}$$

Otherwise, $r_3''(k+1)$ is eliminated, and the above iteration procedure will restart. $r_1(k)$ and $r_2(k)$ are normally randomly selected from the population, and should be mutually different from each other. Apparently, the update of the chromosomes in the DE is similar to the crossover operator of the GA. As a matter of fact, the difference between two chromosomes is an estimation of the gradient information in that zone, where both chromosomes belong to. Therefore, the DE can be considered as a gradient descent-based random search method. Compared with the GA, it has the advantages of algorithm simplicity and optimization efficiency. We apply the DE in re-editing the unqualified NSA detectors so as to reduce the overall cost of detectors generation.

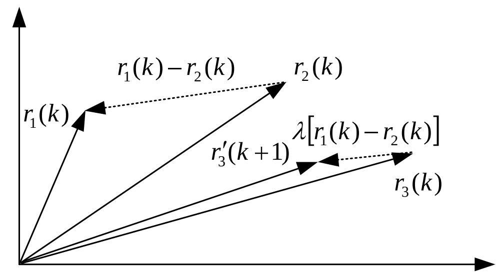


Fig. 2. Principle of Differential Evolution (DE) method.

B. Detectors Re-editing in Negative Selection Algorithm

As discussed above, the unqualified detectors are usually eliminated in the NSA, and new detectors are continuously generated until a given number of detectors are available. Nevertheless, in practice, the generation of detectors could be indeed intensive with regard to both cost and time. Hence, re-editing existing unqualified detectors is sometimes more economical than generating fresh detectors, if the re-editing technique employed is simple and efficient. Our DE-based NSA detectors re-editing scheme is illustrated in Fig. 3. Suppose detector $[w_1, w_2, \cdots, w_L]$ fails to pass the negative selection phase. Two qualified detectors, $[w_1^1, w_2^1, \cdots, w_L^1]$ and $[w_1^2, w_2^2, \cdots, w_L^2]$, are first randomly selected from the detector set. Next, $[w_1, w_2, \cdots, w_L]$ is updated to $[w_1', w_2', \cdots, w_L']$ using the DE method as follows:

$$[w_1', w_2', \cdots, w_L'] = [w_1, w_2, \cdots, w_L] + \lambda\{[w_1^1, w_2^1, \cdots, w_L^1] - [w_1^2, w_2^2, \cdots, w_L^2]\}. \quad (5)$$

$[w_1', w_2', \cdots, w_L']$ is then examined with the self samples again, as in (1) and (2), to check its validity. If $[w_1', w_2', \cdots, w_L']$ is still not qualified, it will be further updated with two newly chosen qualified detectors. In other words, the re-editing of the unqualified detectors is an iterative procedure, which is repeated until $[w_1, w_2, \cdots, w_L]$ become valid or a preset number of DE iterations are reached.

As we know, conventional NSA has the drawback of potential waste of detectors generation resources. The proposed DE-based detectors re-editing system can overcome this shortcoming by re-using the unqualified detectors. Our approach is especially practical in those cases, where it is much more costly to generate new detectors than to modify existing ones. For example, the detectors are usually implemented on electronic circuits in practice. Amending the circuits that have been already designed could be more cost-saving than building new prototypes. Moreover, due to the appropriate computational complexity of the DE method, this novel scheme can result in an accelerated detectors generation process.

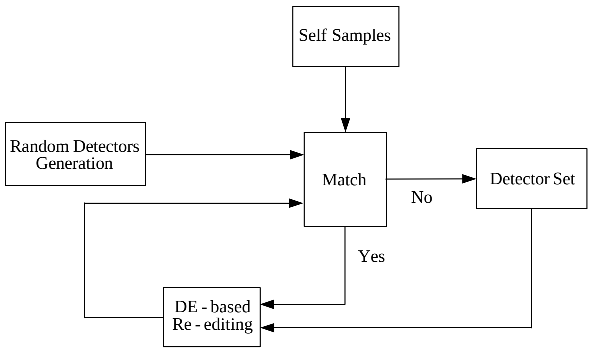


Fig. 3. DE-based detectors re-editing in NSA.

4. Detectors Censoring in Negative Selection Algorithm

It is difficult if not impossible to incorporate domain knowledge of the anomaly to be detected into the NSA detectors generation and selection. However, employment of useful prior information can indeed enhance the anomaly detection performance of the original NSA [7]. In this section, we present a new detectors censoring method, as shown in Fig. 4. The censoring phase is applied to the detector set in order to retain the detectors that are suitable for the specific anomaly detection problems. On the basis of the prior knowledge, the suitability of all the detectors is evaluated, and those inefficient ones are removed from the detector set.

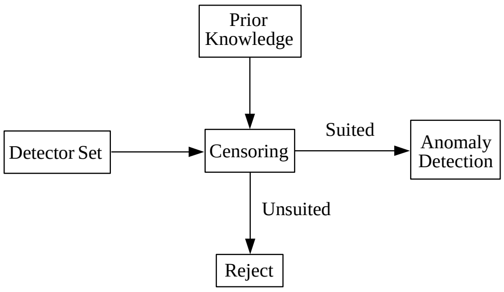


Fig. 4. Detectors censoring in NSA.

There are various ways of utilizing different domain knowledge to censor the NSA detectors. We here only focus on the domain knowledge reflecting the variations of the anomalous signals under detection, because anomaly can often lead to high-frequency oscillations. If known beforehand, the degrees of severity of the signal variations are used for our detectors censoring. More precisely, for a time series signal x_n $(i=1,2,\cdots,n)$, it is split into non-overlapping windows, $[x_1,x_2,\cdots,x_L],[x_{L+1},x_{L+2},\cdots,x_{2L}],\cdots,[x_{n-L+1},x_{n-L+2},\cdots,x_n]$. As an example, the degree of the variation severity of $[x_1,x_2,\cdots,x_L]$, V_1, is calculated with the backward difference technique:

$$V_1=\sum_{i=1}^{L-1}\left|x_{i+1}-x_i\right|. \tag{6}$$

Similarly, $V_2,V_3,\cdots,V_{\frac{n}{L}-1}$ are obtained. Note that as the prior knowledge, the range of $V_1,V_2,\cdots,V_{\frac{n}{L}-1}$ is assumed available in advance. The suitability of all the detectors in the detector set can be evaluated according to (6). For instance, the suitability of $[w_1,w_2,\cdots,w_L]$ is

$$W_1=\sum_{i=1}^{L-1}\left|w_{i+1}-w_i\right|. \tag{7}$$

Based on the range of V_i $(i=1,2,\cdots,\frac{n}{L}-1)$, we can select the detectors in the following way: if W_i of $[w_{(i-1)L+1},w_{(i-1)L},\cdots,w_{iL}]$ is beyond $[\min(V_i),\max(V_i)]$, this detector is expunged from the detector set. Every detector needs to go through the above suitability evaluation and censoring stages. Obviously, the detector set is further tailored to target at dealing with the specific anomaly detection of x_n. In summary, our detectors censoring approach can utilize the prior knowledge of the anomalous signals to provide us with goal-directed detectors. Nevertheless, it has the drawback of demanding for more detectors to be generated, because a certain portion of the detectors are removed from the detector set in the censoring phase. That is to say, this censoring technique actually slows down the detectors generation procedure.

5. SIMULATIONS

In this section, we use three numerical examples to demonstrate the effectiveness of the proposed NSA detectors re-editing and censoring schemes. The first two examples are on the basis of only artificial data, and a bearings fault detection problem is investigated in the third example.

Example 1. DE-based detectors re-editing in negative selection algorithm

In the first example, 1,000 self samples are normalized within [0, 1], and they form the shape of a crossing [14], as represented by '+' in Fig. 5. The radii of all the detectors are chosen to be 0.05. A 100-detector set is first generated using the self samples. Suppose there is an unqualified detector located at (0.5, 0.5), which is denoted by the filled circle. The DE method is next applied to re-edit it. Figure 5 shows a typical evolution procedure, in which only four DE iterations are involved. However, we should point out the number of the iterative steps needed for the detectors re-editing always varies, due to the stochastic nature of the DE technique. This simple example demonstrates that the unqualified detectors can be updated to become qualified in our detectors re-editing scheme.

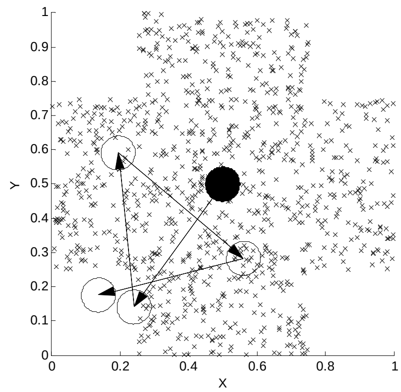


Fig. 5. Example of DE-based detectors re-editing in NSA.

Example 2. Anomaly detection of sinusoidal type signals

The normal and abnormal signals in this example are pure and noise-corrupted sinusoidal type signals with different frequencies, as illustrated in Figs. 6 (a) and (b), respectively. Compared with the normal signal, the abnormal one has a 10-time higher frequency, and it is distorted by white noise. The elevated frequency and noise here are assumed to be caused by the anomaly. Some simulation parameters are given as follows: number of detectors is 100, detector coverage $d = 1$, and width of detectors $L = 10$. Note that these parameters are not guaranteed to achieve the best anomaly detection rate, because they are chosen solely based on *trial and error*. Both the normal and abnormal signals contain 1,000 samples.

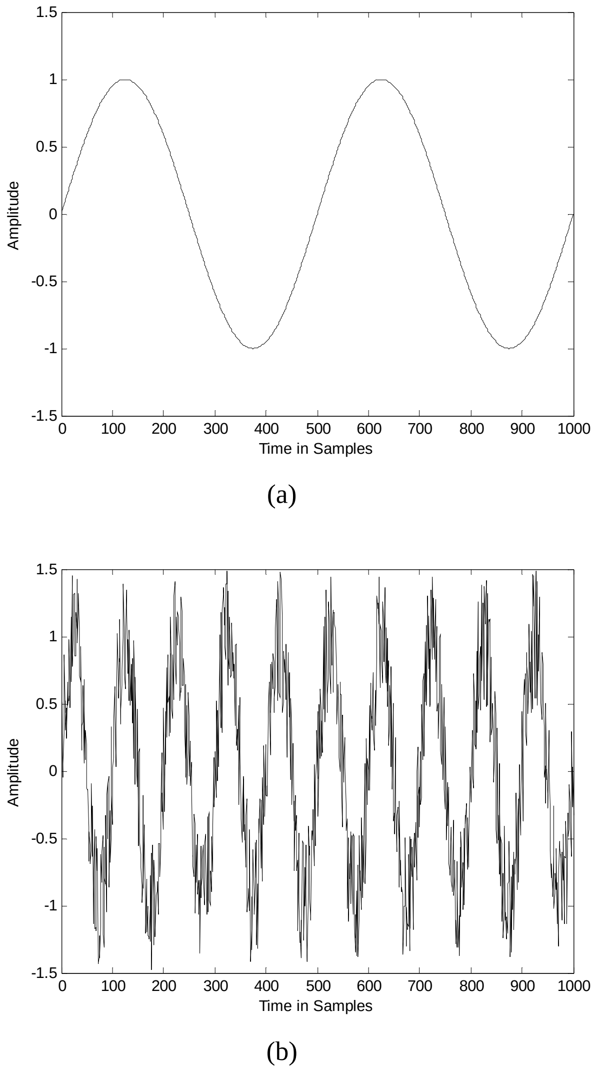


Fig. 6. Sinusoidal type signals in Example 2.
(a) normal signal, (b) abnormal signal.

The degrees of variations of these two signals are measured by V in (6), and are illustrated in Figs. 7 (a) and (b). Apparently, V of the abnormal signal, which is between 1 and 5, is much larger than that of the normal signal. As aforementioned, the range of V is considered as the prior knowledge. Thus, in our detectors censoring system, the suitability of all the detectors in the detector set is evaluated, and only those with the W within [1, 5] can be retained.

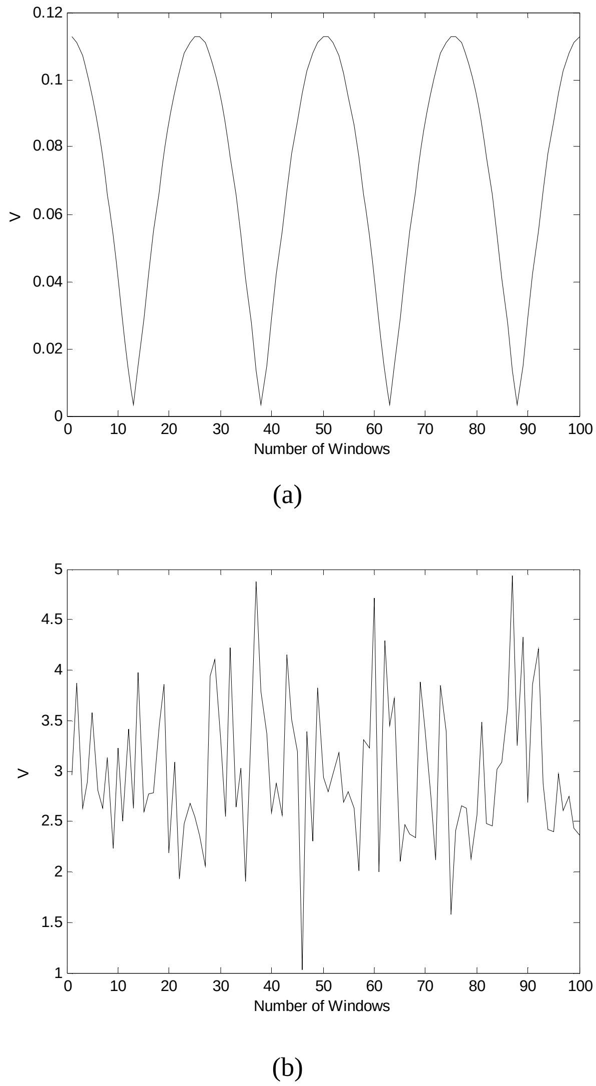
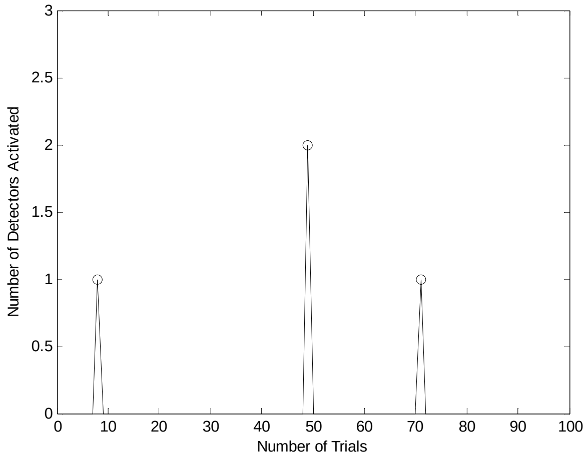
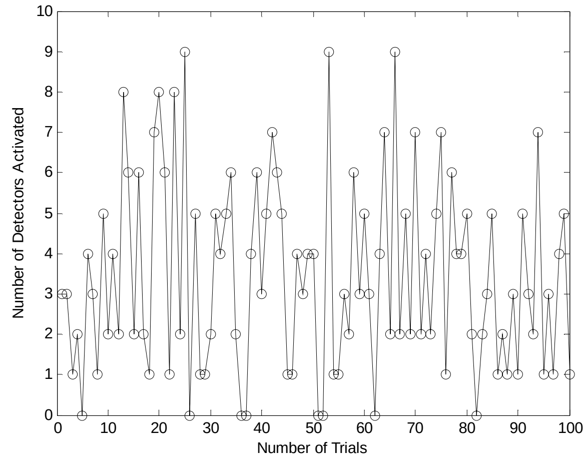


Fig. 7. V of normal and abnormal signals.
(a) V of normal signal, (b) V of abnormal signal.

The anomaly detection results of the detectors before and after the above censoring are demonstrated in Figs. 8 (a) and (b), respectively. The number of the detectors activated by the abnormal signal is deployed to examine their efficiency. We stress that a total of 100 trials are run. For the detectors before censoring, only one or two detectors can detect the anomaly in certain trials. Figure 8 (b) shows that the censored detectors are more efficient for anomaly detection than those in Fig. 8 (a). Averagely, 3.4 detectors are activated in each trial among the ones, which have passed the censoring phase. In other words, a significantly improved anomaly detection performance can be achieved with the detectors censored using the prior information of the anomalous signal.



(a)

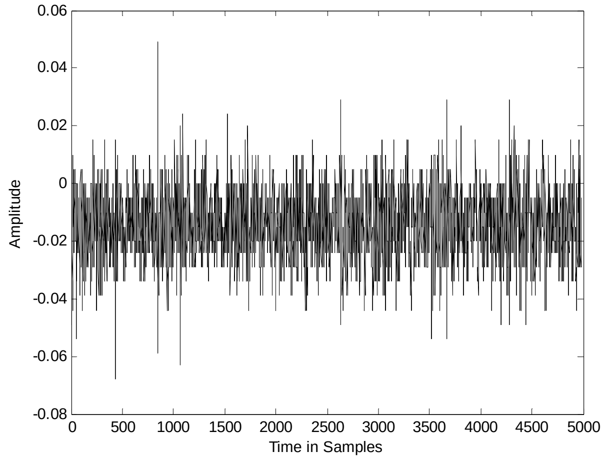


(b)

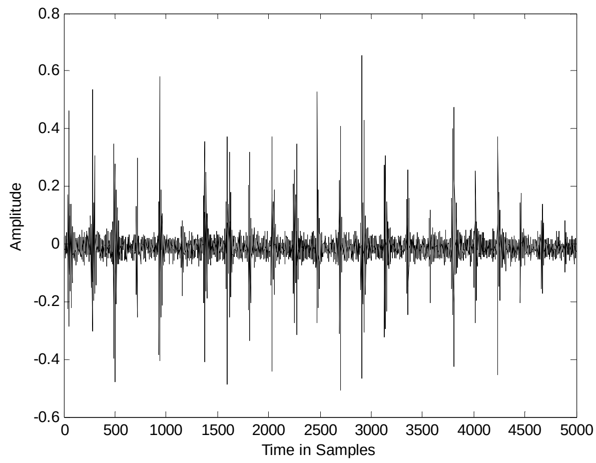
Fig. 8. Anomaly detection results of detectors.
(a) before censoring, (b) after censoring.

Example 3. Bearings fault detection

Bearings are indispensable components in rotating machinery. Therefore, appropriate monitoring of their conditions is crucial to ensure the normal operating status of motors [15]. Because ball damage is a typical bearings fault, our detectors censoring scheme is examined with this fault detection problem. The feature signals of the healthy and faulty bearings are shown in Figs. 9 (a) and (b), respectively. Note that the scales of amplitude in these two figures are different. There are 5,000 samples in both two signals, which are collected at the sampling frequency of 20 kHz from a vibration sensor mounted on top of the NYLA-K eight-ball bearings. The model of the vibration sensor is IMI Sensors 601A01. The motor is a three-phase industrial motor of 0.5 horsepower manufactured by the Baldor Electric Company. It has the rotation speed at 1,782 rpm.



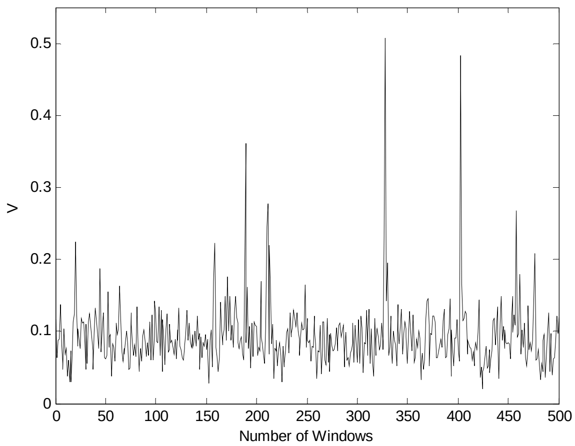
(a)



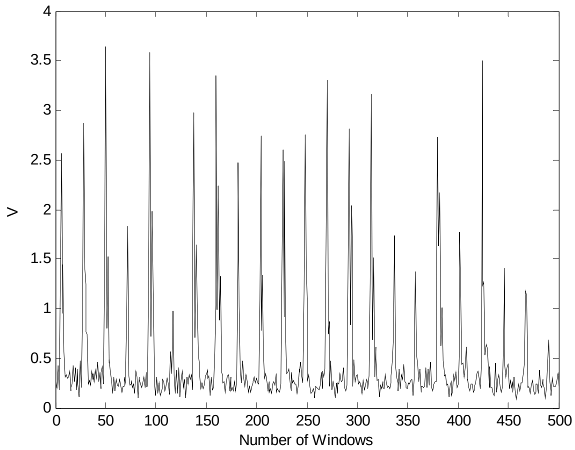
(b)

Fig. 9. Feature signals of bearings in Example 3.
(a) healthy bearings, (b) faulty bearings.

The above two feature signals are split into non-overlapping windows with the width of 10. Their degrees of variations V are given in Figs. 10 (a) and (b). We can observe that due to the existing ball damage, the faulty bearings generate much higher degrees of variations in the feature signal than the normal bearings. We choose the number of detectors to be 1,000, and the detector coverage $d = 0.3$. As in Example 2, the width of detectors $L = 10$. However, the thresholds of W for censoring the detectors are 0.15 and 3 instead of $\min(V_i)$ and $\max(V_i)$. Again, 100 simulation trials have been run. Figures 11 (a) and (b) illustrate the fault detection results of the detectors before and after censoring. The total numbers of the detectors activated by the faulty feature signal are 22 and 71, respectively. It is clearly visible that the censored detectors can yield a superior bearings fault detection performance over the uncensored ones.

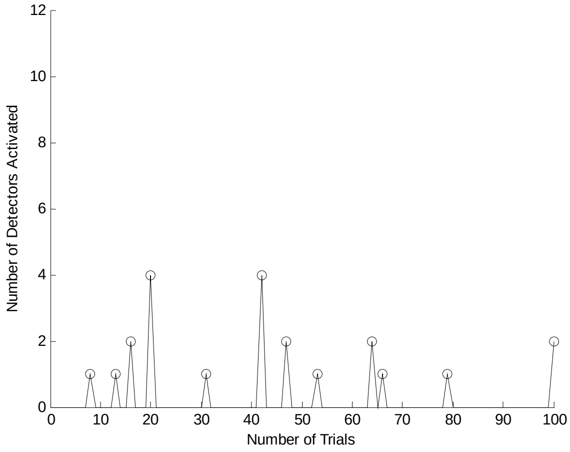


(a)

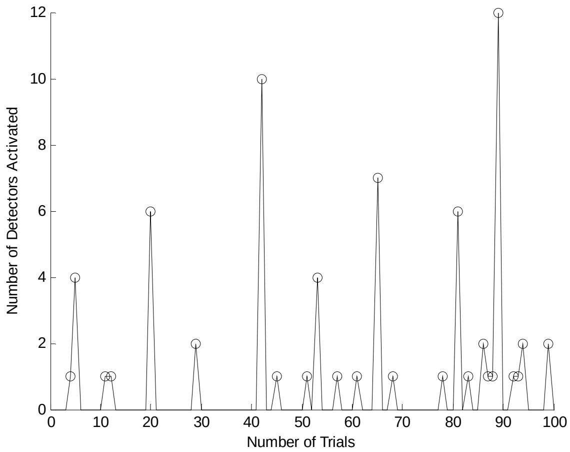


(b)

Fig. 10. V of feature signals of bearings.
(a) healthy bearings, (b) faulty bearings.



(a)



(b)

Fig. 11. Bearings fault detection results of detectors.
(a) before censoring, (b) after censoring.

6. CONCLUSIONS

In this paper, we propose two novel schemes for the NSA detectors re-editing and censoring, in which the unqualified detectors are updated using the DE method to become qualified, and the detector set is censored based on the prior information of the anomaly. Three numerical examples, including a bearings fault detection problem, are employed to verify our approaches. Enhanced performances of anomaly detection are obtained with these schemes in the computer simulations. We emphasize that the domain knowledge is always application dependent, and is not only limited to the severity of variations of the anomalous signals discussed here. The proposed detectors re-editing and censoring techniques can be also generalized to other anomaly and fault detection areas. Therefore, how to apply different kinds of prior knowledge for the detectors censoring remains an interesting research topic.

ACKNOWLEDGMENTS

This research work was funded by the Academy of Finland under Grant 214144.

REFERENCES

[1] G. A. Goldsby, T. J. Kindt, J. Kuby, and B. A. Osborne, *Immunology*. (5^{th} ed.), New York, NY: W. H. Freeman and Company, 2003.

[2] L. N. de Castro and J. Timmis, *Artificial Immune Systems: A New Computational Intelligence Approach*, London, UK: Springer-Verlag, 2002.

[3] D. Dasgupta, "Advances in artificial immune systems," *IEEE Computational Intelligence Magazine*, vol. 1, no. 4, pp. 40-49, November 2006.

[4] S. Forrest, A. S. Perelson, L. Allen, and R. Cherukuri, "Self-nonself discrimination in a computer," in *Proceedings of the IEEE Symposium on Research in Security and Privacy*, Los Alamos, CA, May 1994, pp. 202-212.

[5] D. Dasgupta and F. González, "An immunity-based technique to characterize intrusions in computer networks," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 3, pp. 281-291, June 2002.

[6] D. Dasgupta and S. Forrest, "Tool breakage detection in milling operations using a negative selection algorithm," Technical Report CS95-5, Department of Computer Science, University of New Mexico, Albuquerque, NM, 1995.

[7] Z. Ji and D. Dasgupta, "Revisiting negative selection algorithms," *Evolutionary Computation*, vol. 15, no. 2, pp. 223-251, 2007.

[8] X. Z. Gao, S. J. Ovaska, X. Wang, and M.-Y. Chow, "A neural networks-based negative selection algorithm in fault diagnosis," *Neural Computing & Applications*, vol. 17, no. 1, pp. 91-98, January 2008.

[9] X. Z. Gao, S. J. Ovaska, and X. Wang, "A GA-based negative selection algorithm," *International Journal of Innovative Computing, Information and Control*, vol. 4, no. 4, pp. 971-979, April 2008.

[10] X. Z. Gao, S. J. Ovaska, and X. Wang, "Particle swarm optimization of detectors in negative selection algorithm," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, Montreal, Quebec, Canada, October 2007, pp. 1236-1242.

[11] F. González, D. Dasgupta, and L. F. Nino, "A randomized real-value negative selection algorithm," in *Proceedings of the 2^{nd} International Conference on Artificial Immune Systems*, Edinburgh, UK, September 2003, pp. 261-272.

[12] R. Storn and K. Price, "Differential evolution: A simple and efficient adaptive scheme for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341-359, December 1997.

[13] R. Poli and W. B. Langdon, *Foundations of Genetic Programming*, Berlin, Germany: Springer-Verlag, 2002.

[14] Z. Ji, *Negative Selection Algorithms: from the Thymus to V-detector*, Ph.D. Thesis of Department of Computer Science, University of Memphis, Memphis, TN, 2006.

[15] X. Z. Gao, S. J. Ovaska, X. Wang, and M.-Y. Chow, "A hierarchical optimization scheme for negative selection algorithm detectors in motor fault detection," in *Proceedings of the IEEE Three-Rivers Workshop on Soft Computing in Industrial Applications*, Passau, Germany, August 2007, pp. 163-169.

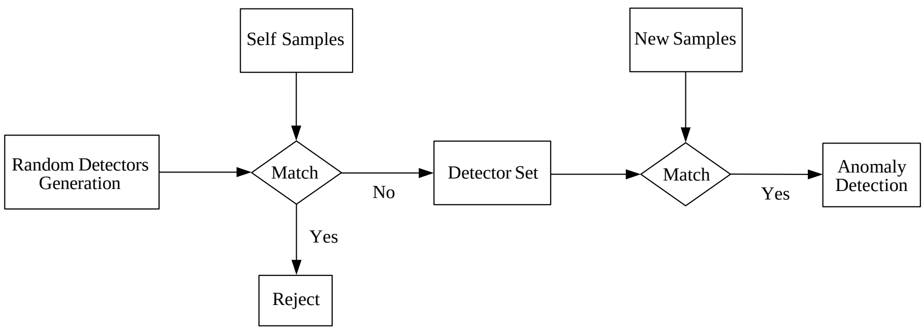


Fig. 1. Negative Selection Algorithm (NSA).