International Journal of Computational Intelligence Systems, Vol.2, No. 3 (October, 2009), 298-311

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

Received: 19/03/09
Accepted: 07/07/09

Abstract

The Negative Selection Algorithm (NSA) is a kind of novelty 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 data 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.

Keywords: artificial immune systems, negative selection algorithm, differential evolution, anomaly detection, fault detection.

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 the 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]. A comprehensive theoretical analysis of the NSA is made in [7], and it is found to have several drawbacks [8]. As we know, the NSA detectors are first generated in a random manner, and undergo the so-called 'negative selection' process thereafter. Only the detec-

tors that do not match the *self* are selected for the anomaly detection, and those unqualified ones will be eliminated. However, practical generation and implementation of the detectors can be rather costly. Therefore, how to reuse the unqualified detectors that are already generated is an important issue. Another shortcoming of the original NSA is that it is difficult if not impossible to explicitly embed the prior information of the novelty to be detected into the detectors selection phase. In this study, we present a Differential Evolution (DE)-based detectors re-editing scheme. A novel method of utilizing the characteristics of the anomalous data for censoring the NSA detectors is also explored.

The remainder of this paper is organized as follows. We introduce the essential principle 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 the 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 cells 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 $[x_1, x_2, \cdots, x_L]$ and $[w_1, w_2, \cdots, w_L]$ be two real-valued vectors denoting a self sample and a candidate detector, respectively, where L is their common order. The matching degree d between $[x_1, x_2, \cdots, x_L]$ and $[w_1, w_2, \cdots, w_L]$ can be calculated based on the Euclidean distance:

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

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 $[w_1, w_2, \cdots, w_L]$ fails to match self sample $[x_1, x_2, \cdots, x_L]$. If $[w_1, w_2, \cdots, w_L]$ 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 $[w_1, w_2, \cdots, w_L]$ matches $[x_1, x_2, \cdots, x_L]$, 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/novelty in the incoming samples. That is, when a new sample $[x_1', x_2', \cdots, x_L']$ matches $[w_1, w_2, \cdots, w_L]$, the existing anomaly is detected.

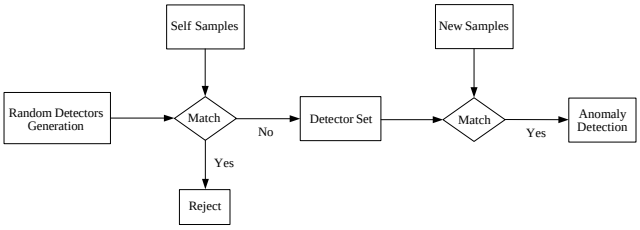


Fig. 1. Negative Selection Algorithm (NSA).

Various types of the real-valued NSA have been studied so far. For example, Ji and Dasgupta propose a variant NSA on the basis of variable-sized detectors (*V*-detectors) [9]. In the *V*-detectors NSA, each self sample has a vicinity (self radius). A *V*-detector is randomly positioned, and its radius is dynamically changed until it reaches the margin of the nearest self sample. Stibor *et al.* compare the *V*-detectors NSA with the Bayesian classification method and one-class Support Vector Machine (SVM) in the anomaly detection [10]. The performance of the *V*-detectors NSA is discovered to be sensitive to several parameters, such as the self radius.

Different from the above real-valued NSA, there are other shape-spaces and matching rules-based NSA [2]. The *r*-contiguous [11] and *r*-chunk [12] matching rules are usually used in the Hamming shape-space, where the self/nonself samples and detectors are represented by binary/character strings. For the *r*-contiguous matching rule, a sample and a detector match, if at least *r* contiguous bits/characters of them are identical. For example, detector [0 *1 1 0* 1 1] matches sample [1 *1 1 0* 0 1], if *r*=3. The *r*-chunk matching rule is actually a generalization of the *r*-contiguous matching rule, which works as follows: a sample and a detector match, if a position exists, from where all the bits/characters of these two are identical over a sequence length *r*. The *r*-chunk matching rule can achieve a better matching performance than that of the *r*-contiguous matching rule [12]. However, both of them cause undetectable elements ('holes') in the Hamming shape-space. The 'holes' are the self samples not available in the detectors generation phase. As a matter of fact, the generation of the *r*-contiguous detectors can be linked to the well-known *k*-CNF satisfiability problem [13]. It has been proved that the Hamming shape-space and *r*-chunk matching rule are only appropriate for the anomaly detection in the low-dimension cases [8].

Conventional NSA has the shortcoming of inefficiency in the detectors generation [14]. That is to say, a lot of randomly generated detectors need to be discarded before the required number of suitable ones are obtained [8]. Several modified versions of the NSA have been investigated during the recent years [15]-[18]. Nevertheless, 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 data. We propose the following detectors re-editing and censoring schemes for 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 [19]. The principle of the DE is similar to that of other evolutionary programming strategies, such as the Genetic Algorithms (GA) [20]. 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 that 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:

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

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 ran-

domly 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:

$$r_3(k+1) = r_3''(k+1). \quad (4)$$

Otherwise, $r_3''(k+1)$ is eliminated, and the above iteration procedure will restart. $r_1(k)$ and $r_2(k)$ are usually 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. The difference between two chromosomes is an estimation of the gradient information in that zone, where both chromosomes belong to. 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. Therefore, we apply the DE in re-editing the unqualified NSA detectors so as to reduce the overall cost of detectors generation.

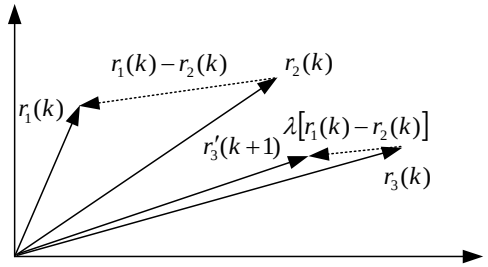


Fig. 2. Principle of Differential Evolution (DE) method.

B. Detectors Re-editing in Negative Selection Algorithm

As discussed above, the unqualified detectors are always eliminated in the NSA, and new detectors are continuously generated until a given number of detectors are available. Unfortunately, in practice, the generation of detectors could be 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. To simplify the presentation in this paper, we only focus on the real-valued NSA.

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:

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

$[w_1', w_2', \cdots, w_L']$ 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 the DE iterations are reached.

As we know, conventional NSA has the drawback of potential waste of detectors generation resources. That is, the randomly generated detectors that do not pass the above negative selection procedure are just disregarded. 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, if the detectors are implemented on electronic circuits, amending the circuits that have been already designed rather than building new prototypes could be cost-saving. Moreover, due to the efficient search capability and moderate computational complexity of the DE method, this novel re-editing scheme can result in an accelerated detectors generation process. To obtain the same number of qualified detectors, it may take less time for the DE-based detectors re-editing scheme than the regular random detectors generation method. This advantage will be demonstrated using computer simulations in the next section.

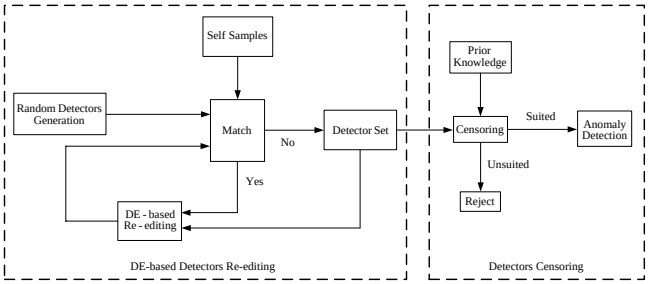


Fig. 3. Detectors re-editing and censoring in NSA.

4. Detectors Censoring in Negative Selection Algorithm

The real-valued NSA requires the nonself examples to achieve a high classification accuracy [8]. Nevertheless, it is difficult if not impossible to incorporate domain knowledge of the anomaly to be detected into the NSA detectors generation and selection. Employment of useful prior information can indeed enhance the novelty detection performance of the original NSA [14]. In this section, we present a new detectors censoring method, as shown in Fig. 3. The censoring phase is applied to the detector set to retain the detectors that are more 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.

There are a lot of ways of utilizing different domain knowledge to censor the NSA detectors. We here only concentrate on the domain knowledge reflecting the variations of the anomalous signals under detection, because anomaly can often lead to high-frequency oscillations. Assumed known beforehand, the degrees of the variation severity of signals 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 ranges of $V_1,V_2,\cdots,V_{\frac{n}{L}-1}$ are 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 ranges of V_i $(i=1,2,\cdots,\frac{n}{L}-1)$, we can select the suited 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 whole detector set is further tailored to target at dealing with the anomaly detection of x_n. In summary, our detectors censoring approach can utilize the prior knowledge of the anomalous signals to provide us with the goal-directed detectors. Nevertheless, it has the disadvantage 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 may slow down the detectors generation procedure.

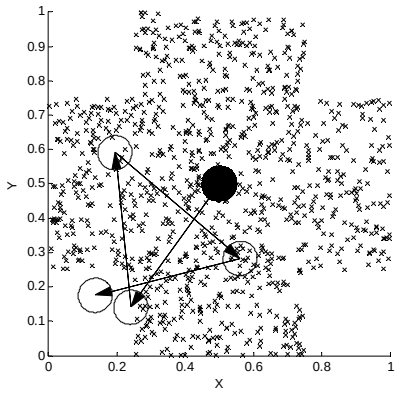
5. Simulations

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, but a practical bearings fault detection problem is investigated in the third example.

Example 1. DE-based detectors re-editing in NSA

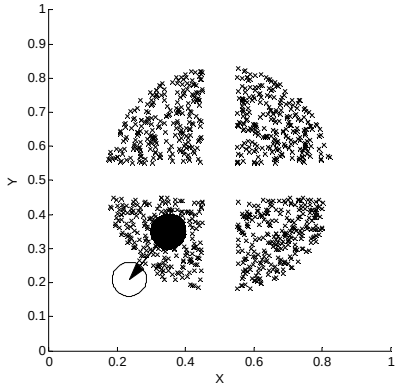
In the first example, two self sample sets, A and B, are

used, each of which consists of 1,000 self samples normalized within [0, 1] [21]. They are represented by '+' in Figs. 4 and 5, respectively. The radiuses of all the detectors are chosen to be 0.05. A 100-detector set is first generated using the self samples. The DE method is next applied to re-edit those unqualified detectors. For self sample set A, the only unqualified detector is located at (0.5, 0.5) denoted by a filled circle. For self sample set B, there are two unqualified detectors located at (0.35, 0.35) and (0.65, 0.65), respectively. Figures 4 (a) and (b) show two typical DE evolution procedures in case of self sample set A, in which three and four iterations are involved, respectively. In Figs. 5 (a) and (b), it takes one and three DE iteration steps to accomplish the detectors re-editing work for self sample set B. However, we must point out that the numbers of the iterative steps needed always vary, due to the locations of the unqualified detectors as well as stochastic nature of the DE technique. Therefore, a total of 10,000 independent trials are made to examine its statistic characteristics. For the three unqualified detectors in Figs. 4, 5 (a), and 5 (b), the average numbers of the DE iterations used for qualifying the detectors are given in Table 1. This simple example demonstrates that the unqualified detectors can be updated to become qualified in our detectors re-editing scheme.

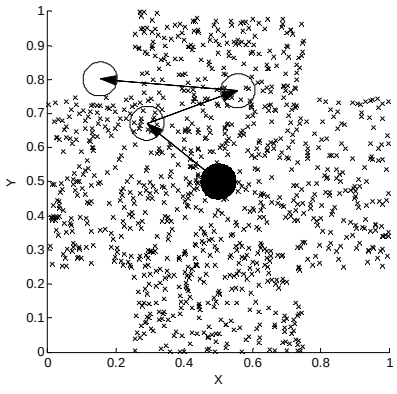


(b)

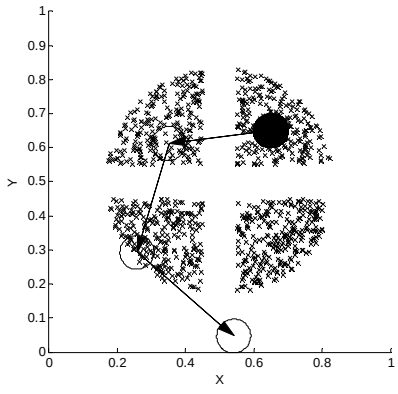
Fig. 4. Examples of DE-based detectors re-editing in NSA (self sample set A).



(a)



(a)



(b)

Fig. 5. Examples of DE-based detectors re-editing in NSA (self sample set B).

Table 1. Average numbers of DE iterations for update of unqualified detectors.

	Average numbers of DE iterations
Unqualified detector in Fig. 5	4.9051
Unqualified detector in Fig. 6 (a)	1.8458
Unqualified detector in Fig. 6 (b)	2.0198

We further compare the efficiency of our DE-based detectors re-editing method and random detectors generation approach. In the former method, the number of the DE iterations for updating 10,000 random unqualified detectors to be qualified is recorded, while in the later approach, we count the total number of the detectors randomly generated after 10,000 qualified ones are obtained. The comparison results are given in Table 2, which are also based on the average of 100 separate trials. The number of the DE iterations used in our method is on the comparable level with that of the detectors generated in the random detectors generation approach. Thus, it can be concluded that the former is a better choice for the NSA than the latter, if the cost of building the detectors is high.

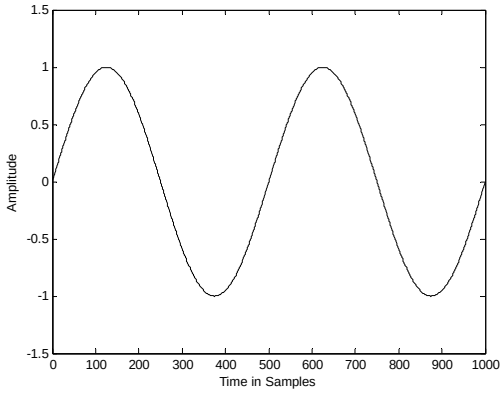
Table 2. Comparison between DE-based detectors re-editing method and random detectors generation approach.
Method 1: DE-based detectors re-editing method (numbers of DE iterations used).
Method 2: Random detectors generation approach (numbers of random detectors generated).

	Method 1	Method 2
Sample set A	48,170	55,111
Sample set B	18,418	17,186

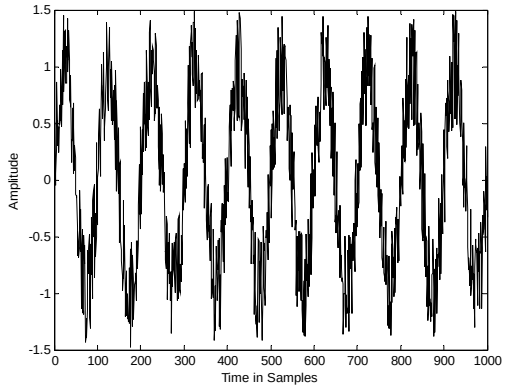
Example 2. Anomaly detection of sinusoidal type signals

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 ten-time higher frequency, and it is distorted by white noise. The elevated frequency and noise here are assumed to be caused by only 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.

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 for the detectors censoring. 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.

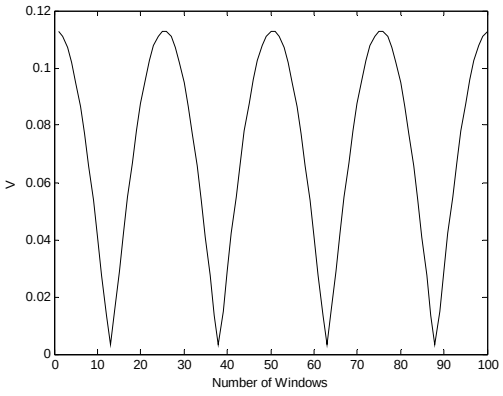


(a)

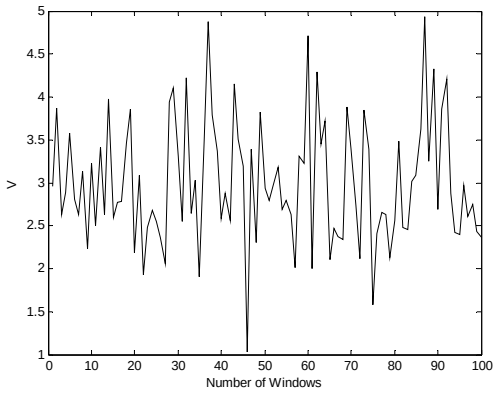


(b)

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



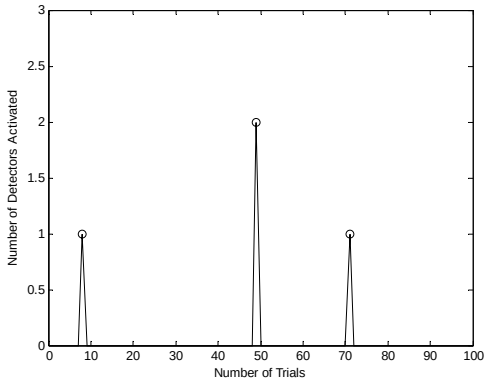
(a)



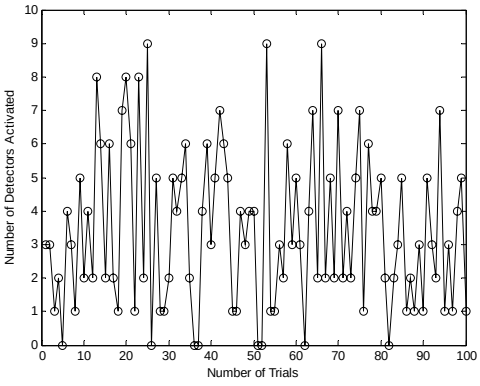
(b)

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 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) anomaly detection results of detectors before censoring,
(b) anomaly detection results of detectors after censoring.

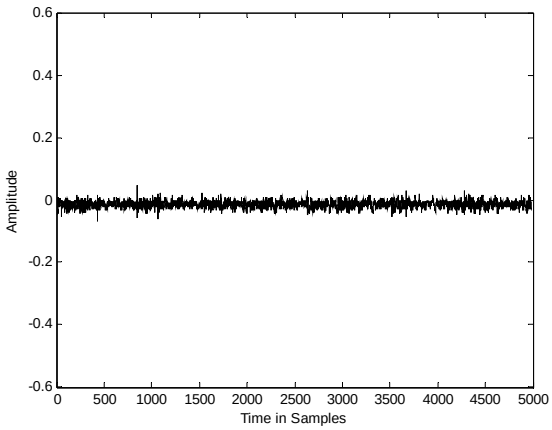
Example 3. Bearings fault detection

Bearings are indispensable components in the rotating machinery. Appropriate monitoring of their conditions is important in maintaining the normal status of operating motors [22]. Unfortunately, various bearings faults, such as ball damage, can occur in practice, due to the severe working environments. An illustrative example of the bearings fault is given in Fig. 9. Two typical kinds of bearings faults, ball damage fault and inner raceway fault, are investigated in our simulations. However, we must emphasize that this paper does not aim at constructing any practical bearings fault detection systems. The bearings fault detection problem is employed here only as a simplified testbed, and most of its technical details are not considered.

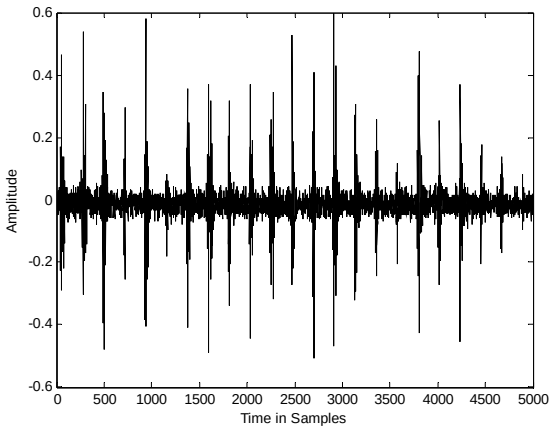
Fig. 9. An illustrative example of bearings fault.

(a) Bearings ball damage fault detection

The feature signals of the healthy and faulty bearings with ball damage are shown in Figs. 10 (a) and (b), respectively. 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 [15]. 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.



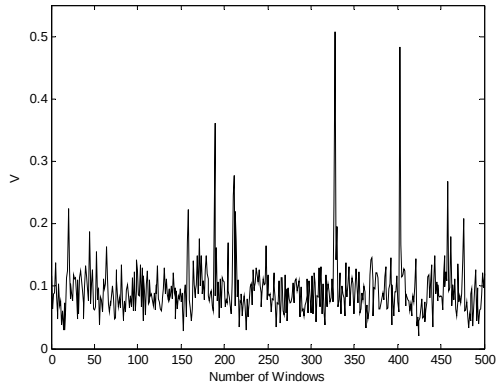
(a)



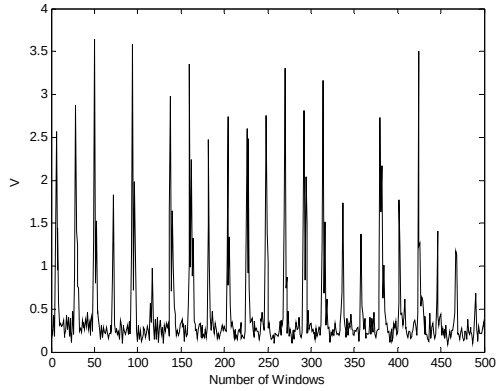
(b)

Fig. 10. Feature signals of bearings in Example 3 (a).
(a) feature signal of healthy bearings,
(b) feature signal of faulty bearings with ball damage fault.

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. 11 (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.



(a)



(b)

Fig. 11. V of feature signals of bearings in Example 3 (a).
(a) V of feature signal of healthy bearings,
(b) V of feature signal of faulty bearings with ball damage.

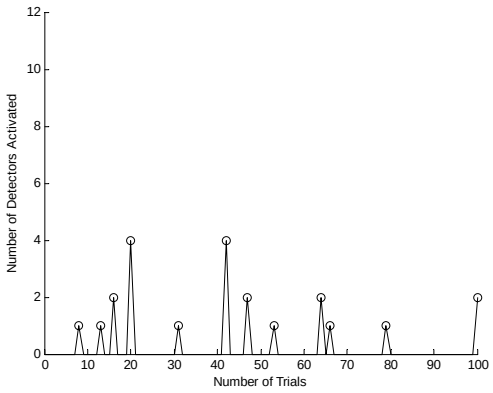
We choose the number of the detectors to be 1,000, and the detector coverage $d = 0.3$. As in Example 2, the width of the detectors $L = 10$. However, the thresholds of W for censoring the detectors are 0.15 and 3 instead of $\min(V_i)$ and $\max(V_i)$. Again, 100 simulation trials have been run. Figures 12 (a) and (b) illustrate the ball damage fault detection results of our detectors. The total numbers of the detectors activated by the faulty feature signal before and after censoring are 22 and 71, respectively. The Received Operating Characteristic (ROC) method is usually used to evaluate the fault detection performance of the NSA [8]. The fault detection and false alarm rates are defined as follows:

$$\text{Detection rate} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{8}$$

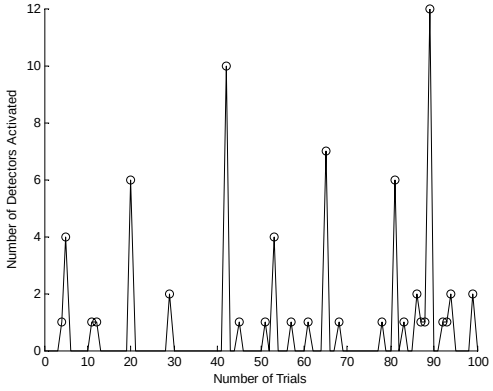
where TP is True Positive, and FN is False Negative.

$$\text{False alarm rate} = \frac{\text{FP}}{\text{FP} + \text{TN}}, \tag{9}$$

where FP is False Positive, and TN is True Negative.



(a)



(b)

Fig. 12. Bearings ball damage fault detection results of detectors.
(a) Bearings ball damage fault detection results of detectors before censoring,
(b) Bearings ball damage fault detection results of detectors after censoring.

The means of the fault detection rates of the detectors before and after censoring are 0.054% and 0.214%, re-

spectively. Their false alarm rates are both zero. The above fault detection results are summarized in Table 3. It is clearly visible that the censored detectors can yield a superior bearings ball damage fault detection performance over the uncensored ones.

Table 3. Bearings ball damage fault detection performances of detectors before and after censoring.

	Detectors before censoring	Detectors after censoring
Numbers of detectors activated	22	71
Detection rates (mean)	5.4×10^{-4}	0.0021
Detection rates (standard derivation)	0.0014	0.0047
False alarm rates (mean)	0	0
False alarm rates (standard derivation)	0	0

(b) Bearings inner raceway fault detection

In the bearings inner raceway fault detection, the motor system consisting of a two horse-power reliance electric motor (left), a torque transducer/encoder (center), a dynamometer (right), and control electronics (not shown) is deployed [23], as illustrated in Fig. 13. The motor speed and motor load are 1,750 rpm and two horse-powers, respectively. The inner raceway fault with a diameter of $0.014''$ is centered at the load zone. The experimental data is collected at 48,000 samples/second from the drive end bearings using the accelerometers attached to the housing with magnetic bases. The vibration signals of the bearings are measured and recorded in a 16 channel DAT recorder. Figures 14 (a) and (b) show those vibration samples from the healthy bearings and faulty bearings with the introduced inner raceway fault, respectively. The degrees of variations V of the feature signals are given in Figs. 15 (a) and (b). More relevant details of the bearings data in our simulations can be found from [23].

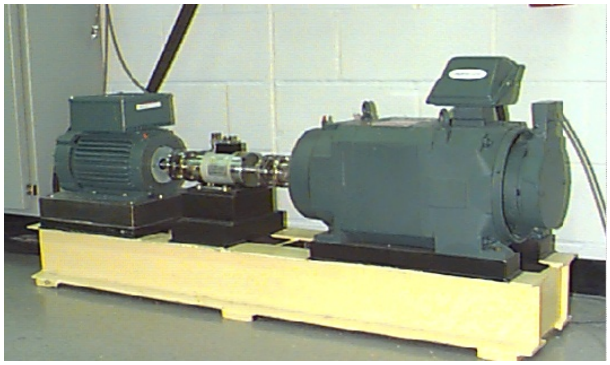
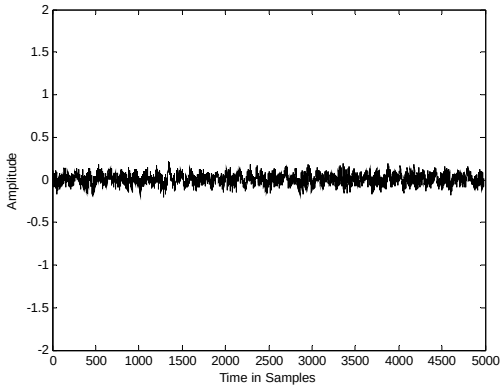
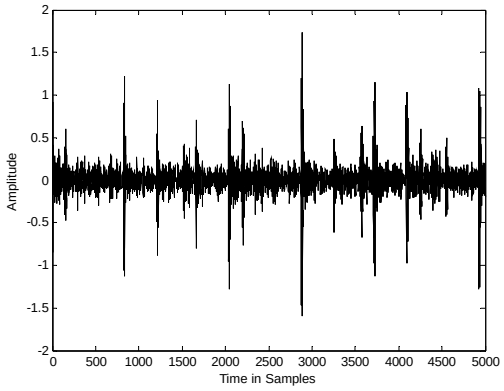


Fig. 13. Motor system in bearings inner raceway fault detection [23].

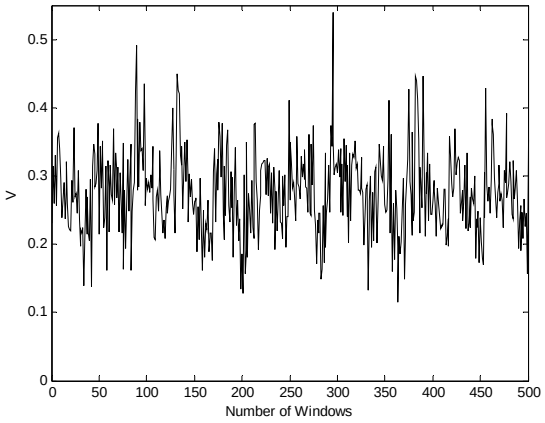


(a)

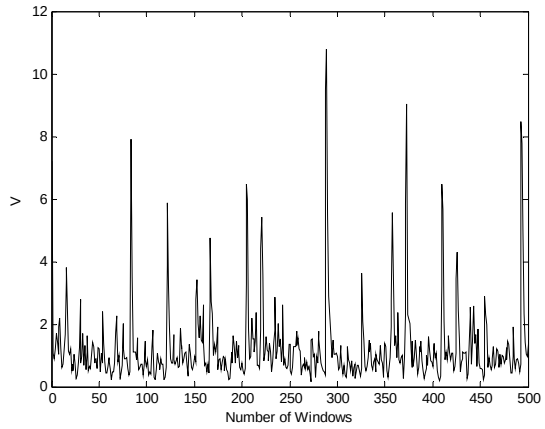


(b)

Fig. 14. Feature signals of bearings in Example 3 (b).
(a) feature signal of healthy bearings,
(b) feature signal of faulty bearings with inner raceway fault.



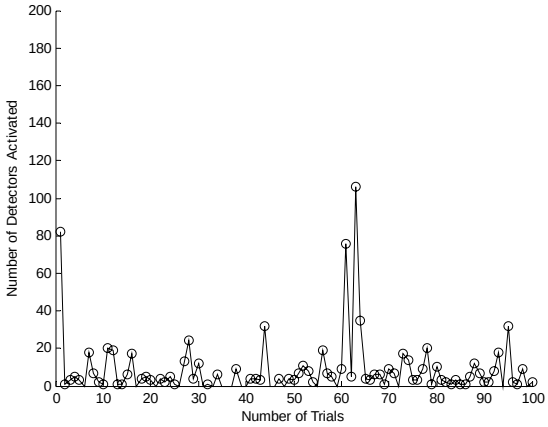
(a)



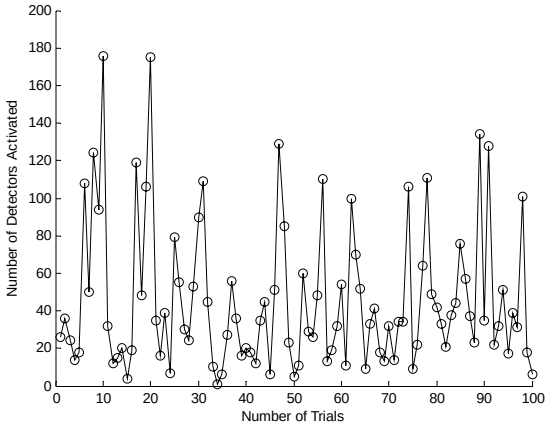
(b)

Fig. 15. V of feature signals of bearings in Example 3 (b).
(a) V of feature signal of healthy bearings,
(b) V of feature signal of faulty bearings with inner raceway fault.

The number of our detectors in the inner raceway fault detection is 100. We choose the detector coverage $d = 2$, and the width of detectors $L = 10$. The two thresholds of W for censoring these detectors are 1 and 6.5. Figures 16 (a) and (b) illustrate the inner raceway fault detection results of the detectors before and after censoring: 4,592 *vs.* 847. The fault detection performances of the detectors in Example 3 (b) are summarized in Table 4, from which we can observe that the proposed censoring approach has significantly enhanced the inner raceway fault detection capability of the NSA detectors.



(a)



(b)

Fig. 16. Bearings inner raceway fault detection results of detectors.
(a) Bearings inner raceway fault detection results of detectors before censoring,
(b) Bearings inner raceway fault detection results of detectors after censoring.

Table 4. Bearings inner raceway fault detection performances

of detectors before and after censoring.

	Detectors before censoring	Detectors after censoring
Numbers of detectors activated	847	4,592
Detection rates (mean)	0.0087	0.0709
Detection rates (standard derivation)	0.011	0.0543
False alarm rates (mean)	6.0×10^{-6}	3.8×10^{-4}
False alarm rates (standard derivation)	4.5×10^{-4}	9.2×10^{-4}
$\frac{\text{Detection rate}}{\text{False alarm rate}}$	145.3	186.6

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 whole detector set is censored based on the prior information of the anomaly. Three numerical examples, including a practical bearings fault detection problem, are employed to verify our approaches. Improved detectors generation and anomaly/fault detection performances are obtained with these two methods in the computer simulations. We emphasize that the domain knowledge used for the detectors censoring is always application dependent, and is not only limited to the variation severity of the anomalous signals discussed here. In addition, study of the robustness of the proposed schemes is an interesting research topic.

Acknowledgments

This research work was funded by the Academy of Finland under Grants 214144 and 124721. The authors would like to thank the anonymous reviewers for their insightful comments and constructive suggestions that have improved the paper.

References

1. G. A. Goldsby, T. J. Kindt, J. Kuby, and B. A. Osborne, *Immunology*. (5th 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. F. Esponda, S. Forrest, and P. Helman, "A formal framework for positive and negative detection schemes," *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics*, vol. 34, no. 1, pp. 357-373, 2004.
8. T. Stibor, P. Mohr, J. Timmis, and C. Eckert, "Is negative selection appropriate for anomaly detection?" in *Proceedings of the 2005 Conference on Genetic and Evolutionary Computation*, Washington DC, June 2005, pp. 321-328.
9. Z. Ji and D. Dasgupta, "Real-valued negative selection algorithm with variable-sized detectors," in *Proceedings of the 2004 Conference on Genetic and Evolutionary Computation*, Seattle, WA, June 2004, pp. 287-298.
10. T. Stibor, J. Timmis, and C. Eckert, "A comparative study of real-valued negative selection to statistical anomaly detection techniques," in *Proceedings of the International Conference on Artificial Immune Systems*, Banff, AL,

Canada, August 2005, pp. 262-275.

11. J. K. Percus, O. E. Percus, and A. S. Perelson, "Predicting the size of the T-cell receptor and antibody combining region from consideration of efficient self-nonself discrimination," *Proceedings of National Academy of Sciences USA*, vol. 90, pp. 1691-1695, 1993.
12. F. González, D. Dasgupta, and G. Gomez, "The effect of binary matching rules in negative selection," in *Proceedings of the 2003 Conference on Genetic and Evolutionary Computation*, Chicago, IL, July 2003, pp. 195-206.
13. T. Stibor, J. Timmis, and C. Eckert, "The link between r-contiguous detectors and k-CNF satisfiability," in *Proceedings of the IEEE Congress on Evolutionary Computation*, Vancouver, BC, Canada, July 2006, pp. 491-498.
14. Z. Ji and D. Dasgupta, "Revisiting negative selection algorithms," *Evolutionary Computation*, vol. 15, no. 2, pp. 223-251, 2007.
15. 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.
16. 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.
17. 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.
18. F. González, D. Dasgupta, and L. F. Nino, "A randomized real-value negative selection algorithm," in *Proceedings of the International Conference on Artificial Immune Systems*, Edinburgh, UK, September 2003, pp. 261-272.
19. 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.
20. R. Poli and W. B. Langdon, *Foundations of Genetic Programming*, Berlin, Germany: Springer-Verlag, 2002.
21. 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.
22. X. Z. Gao, S. J. Ovaska, X. Wang, and M.-Y. Chow, "Multi-level optimization of negative selection algorithm detectors with application in motor fault detection," *Intelligent Automation and Soft Computing* IN PRESS.
23. http://www.eecs.case.edu/laboratory/bearing/welcome_overview.htm.