Chapter 5
Artificial Immune Systems for Anomaly Detection

Eduard Plett
Kansas State University at Salina, USA

Sanjoy Das
Kansas State University, USA

Dapeng Li
Kansas State University, USA

Bijaya K. Panigrahi
Indian Institute of Technology, India

ABSTRACT

This chapter introduces anomaly detection algorithms analogous to methods employed by the vertebrate immune system, with an emphasis on engineering applications. The basic negative selection approach, as well as its major extensions, is introduced. The chapter next proposes a novel scheme to classify all algorithmic extensions of negative selection into three basic classes: self-organization, evolution, and proliferation. In order to illustrate the effectiveness of negative selection based algorithms, one recent algorithm, the proliferating V-detectors method, is taken up for further study. It is applied to a real world anomaly detection problem in engineering, that of automatic testing of bearing machines. As anomaly detection can be considered as a binary classification problem, in order to further show the usefulness of negative selection, this algorithm is then modified to address a four-category problem, namely the classification of power signals based on the type of disturbance.

INTRODUCTION

The vertebrate immune system possesses the ability to recognize, adapt to, and eventually eliminate invasive foreign bodies with remarkable precision. Because of this unique capability, the immune system provides the basis for a number of bio-inspired problem solving approaches in engineering (Castro & Timmis, 2003). These approaches are collectively called *Artificial Immune Systems* (AIS).

DOI: 10.4018/978-1-60566-766-9.ch005

The first stage of the immune system's response to these foreign bodies is their detection. The task of recognizing foreign bodies is done by means of a class of cells called lymphocytes, which are present in the bloodstream. Two kinds of lymphocytes are present, the *B-cells* and the *T-cells*. The B-cells are produced by the bone marrow while the T-cells are generated by a structure called the thymus. The latter cells produce molecules called *antibodies,* which have the ability to bind themselves to specific molecules called *pathogens* that are found in the invasive foreign bodies. Depending on their structure, different antibodies will bind to different types of pathogens, and this ability is called the *affinity* of the antibodies. Further, antibodies must not bind to the molecules produced by their own organism. The ability to distinguish between own cells (called *Self*) and pathogens (called *Nonself*) is termed *Self-Nonself discrimination*. Self-Nonself discrimination is a key feature of the antibodies, and the principles of Self-Nonself discrimination have been successfully applied to many AIS based anomaly detection applications in engineering and computer science (*cf.*Aickelin *et al*. 2004).

Under normal circumstances, Self-Nonself discrimination could work by using *positive* characterization, i.e. train the antibodies to recognize known samples of foreign cells. For example, a typical application of positive characterization is a computer anti-virus program. The virus definitions have to be periodically updated to enable the system to identify new threats. However, it is impossible to predict all possible contaminations by foreign cells, and such a system would be unable to react to threats which is has not encountered before. The vertebrate immune system is, therefore, based on negative characterization or *negative selection* instead (Aickelin *et al*. 2004; Luh & Chen, 2005). This is accomplished by continuously creating a large variety of antibodies. These antibodies are then presented to the body's own cells. If an antibody is found to bind to any of the latter cells, it is simply eliminated from the bloodstream. This is done so that the immune system does not develop an adverse autoimmune reaction. Otherwise, the antibody is released in the bloodstream. The detection process is then straightforward: if an antibody binds to any cell, it is assumed to be foreign and is then destroyed by the immune system.

Artificial immune system algorithms based on negative selection are the mainstay of anomaly detection methods. In a manner reminiscent of their biological counterpart, these algorithms generate a repertoire of *detectors*. These detectors are generated in substantial numbers with the expectation of covering the entire Nonself region. Subsequently, any input that is detected by any detector is classified as an anomalous input.

Negative selection algorithms have been successfully applied to many anomaly detection problems. Probably the most intuitive applications are in enhanced computer security (Dasgupta & Gonzales, 2002; Harmer *et al.*, 2002; Nia *et al.*, 2003). Other applications use negative selection to detect faults in squirrel cage induction motors (Branco *et al.*, 2003), refrigeration systems (Taylor & Corne, 2003), aircraft systems (Dasgupta *et al.*, 2004), and power systems (Gui *et al.*, 2007). They also have been used to detect anomalies in time series data (Nunn & White, 2005) and to recognize patterns, for example the Indian Telugu characters (Ji & Dasgupta, 2004; Ji & Dasgupta, 2006). More recently, a method that uses negative selection in conjunction with an optimization algorithm has been applied to classify faults in rotor rigs from vibration data (Strackeljan & Leiviskä, 2008).

This chapter begins with a description of the basic idea behind negative selection. Next, it focuses on a subclass of detectors, called *V-detectors*, or variable-sized detectors, which is a popular choice in many recent engineering applications. Several recent methods derived from basic negative selection have been outlined, which are divided into three broad categories: self-organizing detectors, evolving detectors and proliferating detectors. In order to demonstrate the effectiveness of negative selection, this chapter describes in detail the proliferation mechanism and proposes extending the detector proliferation