REAL-VALUED NEGATIVE SELETION ALGORITHM FOR ABNORMAL EARTHQUAKE DETECTION

ZEYAD ABD ALGFOOR HASAN

UNIVERSITI TEKNOLOGI MALAYSIA

REAL-VALUED NEGATIVE SELECTION ALGORITHM FOR ABNORMAL EARTHQUAKE DETECTION

ZEYAD ABD ALGFOOR HASAN

A dissertation submitted in fulfilment of the
requirements for the award of the degree of
Master of Science (Computer Science)

Faculty of Computer Science and Information Systems
Universiti Teknologi Malaysia

DECEMBER 2010

To my beloved parents, brothers, sisters and my friends.

ACKNOWLEDGEMENT

All praise be to Allah and may peace and blessings of Allah be upon our prophet, Muhammad and upon all his family and companions .Thanks to Allah who give me good health in my life and thanks to Allah for everything. Without help of Allah, I was not able to achieve anything in this research.

In preparing this thesis, I was in contact with many people, researchers, academicians, and practitioners. They have contributed towards my understanding and thoughts. In particular, I wish to express my sincere appreciation to my thesis supervisor, Prof. Dr. Siti Mariyam Shamsuddin, for encouragement, guidance, critics, advices and supports to complete this research. I really appreciate her ethics and great deal of respect with her students, which is similar to brothers and sisters in the same family.

In addition, I am extremely grateful to my father Dr. Abd Algfoor Hasan for unlimited support and encouragement during this research. My sincere appreciation also extends to Soft Computing Research Group (SCRG) and all my colleagues for the support and incisive comments in making this study a success. Their views and tips are useful indeed. Unfortunately, it is not possible to list all of them in this limited space. However, I must remember that both of Mr. Walid and Nasser for unlimited support. For that, I ask Allah to bless both of them.

ABSTRACT

Earthquake prediction has been a research topic for many years. Many attempts have been made to predict the behavior of earthquake. However, there is yet another field of interest that is seldom explored by the researchers, which is detecting the abnormal behavior of the earthquake. The earthquake magnitude detection studies based on the analysis of historical earthquake data assumes a temporal model. Such models describe the frequencies of occurrence of seismic events as functions of their magnitudes. The most widely used magnitude-frequency model for hazard estimation is that based on the Gutenberg-Richter inverse power law. Artificial Immune System (AIS) has been a common approach in pattern recognition, optimization and many others. However, the application of AIS in the detection of abnormal earthquake behavior is still a new and challenging experience. In this study, Real-Valued Negative Selection Algorithm (RNSA) in AIS is used to establish a model of normal behavior from the large amount of earthquake data and to detect if elements of the data set have changed from an established norm. To show the applicability of the RNSA in abnormal earthquake detection, the earthquake data are divided into several segments and tested according to the assumed normal distribution. Simulation results have revealed that the RNSA improves the performance in terms of detection rate was 87% and 57% for false alarm rate with 8 features.

ABSTRAK

Ramalan gempa bumi telah menjadi topik penyelidikan selama bertahuntahun. Banyak usaha telah dilakukan untuk meramalkan perilaku gempa bumi. Namun, terdapat satu lagi cabangkajian yang jarang diterokai oleh para penyelidik iaitupengesanan perilaku abnormal gempa bumi. Kajian-kajian pengesanan magnitud gempa bumi berdasarkan analisisdata gempa bumi terdahulu membabitkan penggunaan model temporal. Model tersebut menggambarkan frekuensi kejadian peristiwa seismik sebagai suatu fungsi terhadap magnitud-magnitud. Model fungsimagnitud yang digunapakai secara meluas untuk penganggaran bahaya adalah berdasarkan hukum kuasa terbalik Gutenberg-Richter. Sistem kekebalan buatan (AIS) telah menjadi pendekatan lazim di dalam pengecaman corak, pengoptimuman dan banyak lagi. Bagaimanapun, pelaksanaan AIS di dalam pengesanan perilaku abnormal gempa bumi masih lagi baru dan penuh cabaran.Untuk kajian ini, Algoritma Pemilihan Negatif Nilai-Nyata(RNSA) di dalam AIS digunakan untuk membina model perilaku normal daripada sejumlah data gempa bumi bagi mengesan unsur-unsur yang telah berubah berdasarkan model awalan. Bagi menunjukkan keberkesanan RNSA di dalam pengesanan gempa bumi abnormal, data gempa bumi dibahagikan kepada beberapa segmen dan diuji terhadap taburan normal sedia ada. Keputusan Sitimulasi telah menyatakan bahawa RNSA membuktikan pencapaian dalam bentuk kadar kepastian di mana 87% dan 57% adalah kadar pemberitahuan yang salah dengan 8 ciri.

TABLE OF CONTENTS

CHAPTER	TITLE	PAGE
	DECLARATION	ii
	DEDICATION	iii
	ACKNOWLEDGEMENT	iv
	ABSTRACT	v
	ABSTRAK	vi
	TABLE OF CONTENTS	vii
	LIST OF TABLES	x
	LIST OF FIGURES	xi
	LIST OF ABBREVIATIONS	xvii
1	**INTRODUCTION**	**1**
	1.1 Introduction	1
	1.2 Problem Background	3
	1.3 Problem Statement	7
	1.4 Research Aim	8
	1.5 Research Objectives	8
	1.6 Research Scope	9
	1.7 Thesis Organization	9
2	**LITERATURE REVIEW**	**10**
	2.1 Introduction	10
	2.2.1 Body Waves	11

2.2.1.2 Secondary Wave (S-Wave 14

2.2.2 Surface Waves 14

2.2.2.1 Love Waves (LQ) 15

2.2.2.2 Rayleigh Wave (LR) 15

2.2.3 Guided Waves 16

2.2.3.1 Lq Waves 16

2.2.3.2 Rq Waves 17

2.2.3.3 T-Waves 17

2.3 Negative Selection Algorithm 18

2.3.1 Past Research Negative Selection Algorithm in Anomaly detection 21

2.4 Real-valued Negative Selection Algorithm 24

2.5 Related Works on Earthquake Data Based on Computational Intelligence 28

2.5.1 Neural Network 28

2.5.2 Genetic Algorithm 29

2.6 Summary 29

3 RESEARCH METHODOLOGY 30

3.1 Introduction 30

3.2 Literature Study and Review of Existing 32

3.2.1 Determination of Theory, Problem, Objectives and Scope 32

3.2.2 Literature Study and Review of Existing Earthquake Techniques 33

3.3 System Design 33

3.4 System Implementation 33

3.4.1 Data Collection 34

3.4.2 Pre-processing Earthquake Raw Data 37

3.4.3 Training Techniques RNSA, Clonal and BPNN 37

3.5 Performance and Evaluation 38

3.6 Summary 39

4 DESIGN AND IMPLEMENTATION 40

4.1 Introduction 40
4.2 Develop RNSA Classifier for Earthquake 41
4.3 Data Collection 43
4.3.1 Seismicity Indicators 43
4.4 Data Presentation 49
4.5 Real-valued Negative Selection Algorithm 52
4.5.1 Define Self Data 55
4.5.2 Define Candidate Detectors 55
4.5.3 Convert Data to Real-Valued 55
4.5.4 Store as Detectors 56
4.5.5 Presents Test Data 56
4.5.6 Matching Process 57
4.5.7 Notify Detection 57
4.6 Summary 58

5 RESULTS AND DISCUSSION 59
5.1 Introduction 59
5.2 Earthquake Data Selection and Distribution 60
5.2.1 Number of Detectors 61
5.2.1.1 Real-valued Negative Selection Algorithm 62
5.2.1.2 Clonal Selection Algorithm Implementation with Earthquake 67
5.3 Experimental Setting 71
5.4 Result Comparison and Discussion 72
5.6 Summary 75

6 CONCLUSION AND FUTURE WORKS 76
6.1 Introduction 76
6.2 Summary 76
6.3 Contributions of the Study 77
6.4 Recommendations for Future Works 78

REFERENCES 79

LIST OF TABLES

TABLE NO.	TITLE	PAGE
1.1	The previous researchers on earthquake	4
2.1	The Past Research of Negative Selection Algorithm in anomaly detection	22
3.1	The raw data from Northern California Data earthquake center	34
3.2	A list of abbreviations from Northern California website	35
4.1	Eight seismicity indicators	48
4.2	Converting method of normalization data	50
4.3	Converting values from numerical values of each equation into real-valued	51
4.4	Eight seismicity indicators after normalization	52
5.1	Eight seismicity indicator Panakkat and Adeli (2007)	60

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE
1.1	Global map	2
1.2	Seismogram	2
2.1	ground motion for four types of earthquake	12
2.2	Seismogram showing P, PP, S, LQ and LR phases	15
2.3	Seismograms showing *L*g and *R*g. (a) *L*g	16
2.4	Seismogram showing P, S, and T-phases	17
2.5	Pattern Recognition via the Negative Selection Algorithm (a) and (b)	20
2.6	Illustrates an iteration of the RNSA	25
2.7	Real-value negative selection algorithm	26
3.1	Flow chart diagram for research methodology	31
4.1	Northern California website	41
4.2	A framework for detection earthquake based on RNSA	42
4.3	Block diagram of RNSA implementation	54
5.1	Chart RNSA correctly detected with (eight seismicity)	62
5.2	Chart RNSA incorrectly detected with (eight seismicity)	62
5.3	Chart RNSA correctly detected with (T, M_{mean}, $dE^{1/2}$)	63
5.4	Chart RNSA incorrectly detected with (T, M_{mean}, $dE^{1/2}$)	63
5.5	Chart RNSA correctly detected with ($\beta, \eta, \Delta M$)	64
5.6	Chart RNSA incorrectly detected with ($\beta, \eta, \Delta M$)	64
5.7	Chart RNSA correctly detected with	

	(T , M_{mean} , $dE^{1/2}$, β , η , ΔM).	65
5.8	Chart RNSA incorrectly detected with	
	(T , M_{mean} , $dE^{1/2}$, β , η , ΔM).	65
5.9	Chart RNSA correctly detected with (μ,c)	66
5.10	Chart RNSA incorrectly detected with (μ,c)	66
5.11	Chart Clonal correctly detected with (eight seismicity)	67
5.12	Chart Clonal incorrectly detected with (eight seismicity)	67
5.13	Chart Clonal correctly detected with (T, M_{mean} , $dE^{1/2}$)	68
5.14	Chart RNSA incorrectly detected with (T, M_{mean} , $dE^{1/2}$)	68
5.15	Chart Clonal correctly detected with (β , η , ΔM)	69
5.16	Chart Clonal incorrectly detected with (β , η , ΔM)	69
5.17	Chart Clonal correctly detected with	
	(T , M_{mean} , $dE^{1/2}$, β , η , ΔM).	70
5.18	Chart Clonal incorrectly detected with	
	(T , M_{mean} , $dE^{1/2}$, β , η , ΔM).	70
5.19	Chart Clonal correctly detected with (μ,c)	71
5.20	Chart Clonal correctly detected with (μ,c)	71
5.21	Comparison of overall detection rate of three algorithms	73
5.22	Comparison of overall False rate of three algorithms	74

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
AIS	Artificial Immune System
ANN	Artificial Neural Network
BPNN	Back-propagation Neural Network
RNSA	Real-valued Negative Selection Algorithm
NCEDC	Northern California Earthquake Data Centre
GA	Genetic Algorithm
BIS	Biological Immune System
LMBP	feed-forward Levenberg-Marquardt backpropagation
M	Magnitude
RNN	Recurrent Neural Network
RBF	Radial Basis Function
GP	genetic programming
PGA	peak ground acceleration
ANN	A probabilistic neural network
NSIN	Neural Systems identification
P-wave	primary wave
S-wave	Secondary wave
LQ	Love Wave
LR	Rayleigh Wave

CHAPTER 1

INTRODUCTION

1.1 Introduction

Nowadays, Earthquakes are one of the most devastating natural disasters on earth. A strong earthquake is a natural disaster which brings sudden fatality, great economic loss and shock to the community. Earthquakes may occur naturally or because of human activities. The point on the ground surface immediately above the initial rupture point is called the “epicentre” of the earthquake. Usually the earthquake occurs in everywhere, but there are some locations high ratio occurrences more than others. In figure 1.1 shows the global map and black points (epicenter) for the quake-hit its incidence is high. The black points are output for instruments designed to detect and measure vibrations within the earth known as (seismograms).

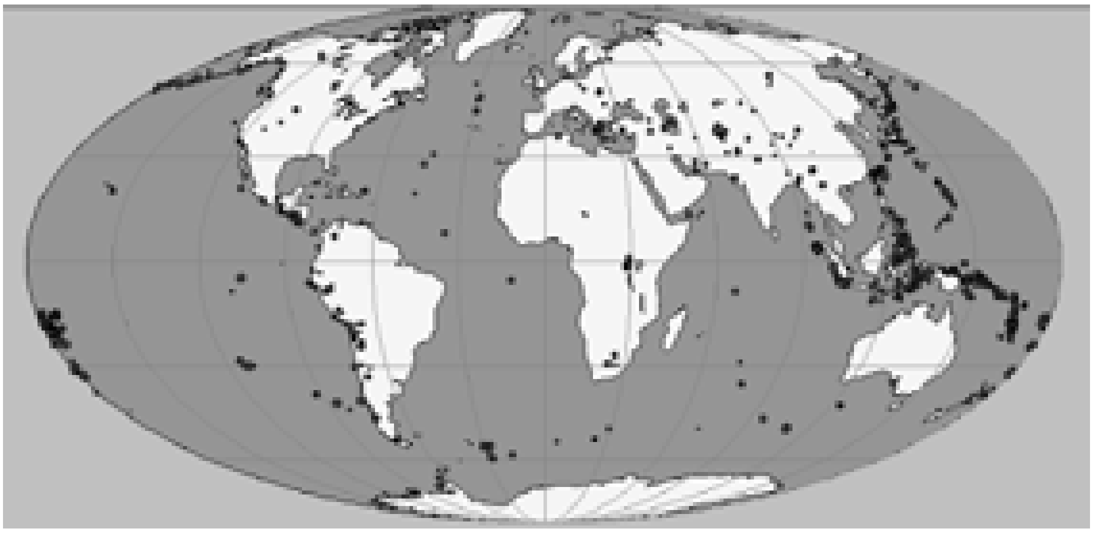

Figure 1.1: The most of earthquakes occur usually around the world in places like California and Alaska in USA, In addition Guatemala. Chile, Peru, Indonesia, Iran, Pakistan, Azores, Portugal, Turkey, New Zealand, Greece, Italy, and Japan, but seismic can occur almost in everywhere.

Seismograms are recordings of ground motion. The ground is continuously at unrest mainly due to waves in the ocean. Sometimes higher amplitude motions are recorded and we talk about a seismic event (see Figure 1.2). Seismic events are caused by a sudden release of energy by seismic sources which are mainly earthquakes, but which also can be explosions, volcanic eruptions, rock-falls etc, Jens Havskov and Lars Ottemöller (2010).

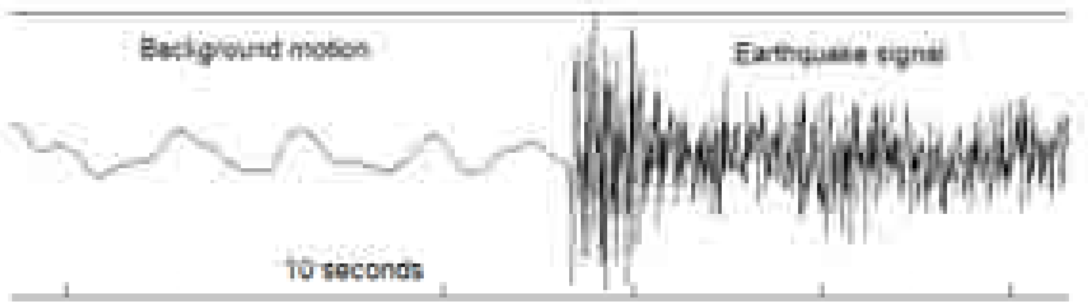

Figure 1.2 Seismogram from a M = 3.8 event in Venezuela. At the left part of the seismogram is seen the natural background noise of the earth and to the right the earthquake signal. The station recording the event is BAUV and the time of the earthquake is 2003 0422 13:029.

Earthquake hazard is greatest disaster in this world. May be what happened in Haiti is very clear example in Tuesday, January 12, 2010, time it took 35 second and the magnitude was 7.0 in Richter measure, it left around 100 thousand killed and $13B according to a study by the Inter-American Development Bank (CNN).

In this study, will be focused on data obtained from Northern California Earthquake Data Centre (NCEDC) in order to detect abnormal behavior for earthquake by using Artificial Immune System (AIS). Real-Valued Negative Selection Algorithm is one of AIS approach will be used to detect the abnormal earthquake. Applying the systems which are biologically inspired such as Neural Networks (NN),Genetic Algorithm (GA) evolutionary computation, DNA computation, natural immune system and so on in the earthquake (seismic wave) has lately attracted a great deal of attention. More recently, considerable research challenges have focused on the exploitation of the key features of the Biological Immune System (BIS) such as recognition, feature extraction, diversity, learning memory, distributed detection, self-regulation and adaptability.

1.2 Problem Background

For seismology, these should be easy. It is hard to imagine topics more interesting than structure and evolution of a planet, as manifested by phenomena as dramatic as earthquake. There are many methods used in seismology, it considers as primary tool for study of the earthquake like physical properties, the existence of the earth's shallow crust, deeper mantle, liquid outer core and solid inner core inferred from variations in seismic velocity with depth (Seth Stein and Michael Wysession, 2003). According to this kind of data, many studies have been appareled to handle with it. Table 1.1 shows studies tried to detect or predict earthquake data.

Table1.1: The previous researchers on earthquake

Technique	researchers	Description	findings
A neural-network model for earthquake occurrence	Bertalan Bodri(2001)	The neural network in this article based on three-layer feed-forward neural network models were constructed to analyze earthquake occurrences and Numerical experiments have been performed with the aim to find the optimum input set configuration which provides the best performance of a neural network	Seismicity rate variations in the Carpathian Pannoman region, Hungary, and the Peloponnesos–Aegean area, Greece, have been used to develop neural network models for the prediction of the origin times of large (M>=6.0) earthquakes.
Neural Network models for earthquake magnitude prediction using multiple seismicity indicators	Ashif Panakkat and Hojjat Adeli (2007)	In this article has used three different neural networks to predict earthquake as fallow:- 1-feed-forward Levenberg-Marquardt backpropagation (LMBP) neural network 2-recurrent neural network(RNN) 3-radial basis function (RBF) neural network	In this article the authors tries to find good technique to predict the earthquake. After test the all of them, the recurrent neural networks was the best one. It have the inherent capacity to model time-series data better compared with other networks.

Genetic Programming-based attenuation relationship: An application of recent earthquakes in turkey	Ali Firat Cabalar and Abdulkadir Cevik (2009)	Applying genetic programming (GP) for the prediction of peak ground acceleration (PGA) using strong-ground-motion data from Turkey. Database has been evaluated by using the best NN models.	The major advantage of GP conventional regression techniques is that there is no predefined function to be considered for modeling. To make sure that GP can be effectively and safely used in modelling earthquake data.
A probabilistic neural network for earthquake magnitude prediction	Hojjat Adeli and Ashif Panakkat (2009)	A probabilistic neural network (PNN) is presented for predicting the magnitude by applying eight computed mathematically parameters known as seismicity indicators	PNN used to predict the earthquake magnitude between 4.5 until 6.0. The PNN based on history record data for seismic events and last probabilistic studies.
Structural damage detection using the optimal weights of the approximating artificial neural networks	Shih-Lin Hung and C. Y. Kao (2002)	In this article presents a novel neural network comprises tow steps. Systems identification (NSIN), structural damage detection (NDDN)	By using two neural networks. First NSIN used to identify the damaged and undamaged states of the structural system. Physical system properties are

			not available in detection phase. Therefore, by supposing some a priori information about system
A neural network approach for structural identification and diagnosis of a building from seismic response data	Huang, Hung and Tu (2003)	This article presents a back-propagation neural network approach. This algorithm trained by using five-story steel frame subjected to different strengths	The results came out between 52% until 60% to diagnose a damage structure. Therefore, in order to detect the location of the damage, this approach needs more to improve or verified

In the previous works, appeared many problems made the works not completely. Like, some techniques work between 4.5-6.0 magnitude, false prediction for long period prediction (must divide the period into less than month) or small regions.

Earthquake prediction is the biggest unsolved problem of seismology. The earthquake needs Long-term predictions, the main idea for detection or prediction depends on the way for pre-processing data. The subject of major interest in the present work, are made a few years to a few decades before the expected earthquakes. They are based generally on analysis of earthquake recurrence times and changes of broad seismicity patterns (Carlson, 1991).

1.3 Problem Statement

Most of the researchers concentrate on the characteristics and potentials for each technique in terms of capability of solve a problem in less time and high efficiency. Therefore, the previous works were focused on use the soft computing scope. The last works based on enhance the techniques and watch the performance for that technique. They are using earthquake data which is considering as time series data.

Some results were somehow satisfactory but not high efficiency. The previous works were focused on three points as followed:-

i. Classification: to classify the magnitude into multi-classes by using threshold. For example, 4.5 - 4.9M, 5-5.4M that means the threshold is (0.5) so on. The target is divided the magnitude into some classes to be easier to handle with.
ii. Detection: to find the abnormal events through seismic data.
iii. Prediction: to predict the earthquake based on pervious studies on seismology.

In this study, the detection of abnormal behavior of earthquake data is the main interest to be studied. As seen in previous section, there are many of studies have tried to find the best solutions for earthquakes. In the past, researches have used algorithms for prediction and detection the earthquake behavior and have clear applications for the use in detect or predict the earthquake. The last studies showed some weaknesses in terms of detection or prediction. However, the new areas of biologically inspired computing using up relatively, there is less research done on application of AIS in physical phenomena. Therefore, there are many models never tried in this area to prove its efficiency. Therefore, the problem statement for this study could be expressed as follow: - How Immune-based solutions could detect the abnormality of earthquake magnitude efficiently?

1.4 Research Aim

The aim of this research to apply artificial intelligence technique RNSA to detect abnormal behavior in earthquake magnitude.

1.5 Research Objective

Intelligent techniques have been widely implemented in magnitude earthquake. However, most of these techniques are employed in detection. Hence, this project is carried out with following objectives:-

i. To develop RNSA Algorithms for magnitude earthquake detection.
ii. To analyze the effectiveness of RNSA in earthquake detection problem.
iii. To compare the results of RNSA for earthquake detection with Clonal Selection and Backpropagation Neural Network (BPNN).

1.6 Research Scope

1. Real-Valued Negative Selection Algorithm is used to establish a model of normal behaviour from the large amount of data and to detect if elements of a set of data have changed from an established norm.

2. The programs are built on windows environment using, Microsoft Office Excel to pre-processing data and Matlab programming languages.

3. Data used for testing and evaluating of the proposed method are obtained from Northern California Earthquake Data Centre (NCEDC). The data include events magnitudes occurred for long period time.

1.7 Thesis Organization

This thesis contains six chapters and organized as follows: - Chapter 1 provides a brief introduction of the study. It covers topics on problem background and motivations, problem statement, research objectives, research scope and thesis organization. Chapter 2 provides the relevant background of Seismic Waves area. Moreover, the relevant artificial intelligence techniques are presented in this chapter, and these include artificial neural networks and artificial immune system and Genetic algorithm. Chapter 3 describes in-depth methodology used in this study. The research methodology is presented as flow chart diagram that describes how each step is carried out. Chapter 4 discusses the design concepts and simulation implementation of RNSA and pre-processing data based on eight mathematical equations. Chapter 5 presents and discusses the experimental results. The performance metrics in terms of detection rate and false alarm rate have been used to analyze and evaluate the effectiveness of the performance of the proposed algorithm. Finally, Chapter 6 concludes the thesis with a summary of the work that has been done and recommendations for future work.

REFERENCES

Adeli, H. and Panakkat, A. (2009). A probabilistic neural network for earthquake magnitude prediction. Neural Networks.

Bodri, B.(2001). A neural-network model for earthquake occurrence. Journal of Geodynamics.

Cabalar, A.F. and Cevik, A.(2009). Genetic programming-based attenuation relationship: An application of recent earthquakes in turkey. Computers & Geosciences.

Carlson, B. and SHAW, JM and Langer, JS. (1992). Patterns of seismic activity preceding large earthquakes. Journal of Geophysical Research.

Dasgupta, D.(1997). Artificial Neural Networks and Artificial Immune Systems: Similarities and Differences. IEEE.

Dasgupta, D.(1999). Immunity-Based Intrusion Detection System: A General Framework. Proc. of the 22nd NISSC.

Dasgupta, D.(2000). An agent based architecture for a computer virus immune system. Proc.GECCO Workshop Artificial Immune System.

Dasgupta, D., Coa,Y., Yang, C.(1999).An immunogenetic Approach to SpectraRecognition . Proc. of the Genetic and Evolutionary Computation Conference.

de Castro, L.N., Timmis, J.(2002). Artificial Immune Systems: A New Computational Intelligence Approach .London: Springer.

Forrest, S., et al. (1997). Computer Immunology. Communications of the ACM, pp 88-96.

Forrest, S., Perelson, A. S., Allen, L. and Cherukuri, R. (1994). Self-Nonself Discrimination in a Computer. *Proceeding of IEEE Symposium on Research in Security and Privacy*, pp. 202 – 212, Oakland.

Gonzalez, F., Dasgupta, D. (2002). A Immunogeneric Approach to Intrusion Detection. GECCO.

Gonzalez, F., Dasgupta, D. and Gomez, J. (2003). The Effect of Binary Matching Rules in Negative Selection. Lecture Notes in Computer Science, vol. 2723, pp. 198-209, Springer-Verlag.

Gonzalez, F.A., Dasgupta, D. (2002). Anomaly Detection Using Real-Valued Negative Selection. Genetic Programming and Evolvable Machine.

Huang, C.S. and Hung, SL and Wen, CM and Tu, TT.(2003). A neural network approach for structural identification and diagnosis of a building from seismic response data. Earthquake Engineering & Structural Dynamics.

Hung, S.L. and Kao, CY.(2002). Structural damage detection using the optimal weights of the approximating artificial neural networks. Earthquake Engineering & Structural Dynamics.

Panakkat, A. and Adeli, H.(2007). Neural network models for earthquake magnitude prediction using multiple seismicity indicators.

Savage. (1999).Seismic anisotropy in the mantle transition zone beneath Fiji-Tonga. Geophys. Res. Lett.