Journal of Informatics and Mathematical Sciences
Vol. 8, No. 4, pp. 225–234, 2016
ISSN 0975-5748 (online); 0974-875X (print)
Published by RGN Publications

http://www.rgnpublications.com

Special Issue
CRIMSTIC 2016
Current Research in Information Technology, Mathematical Sciences, Science and Technology International Conference 2016
April 13-14, 2016, Melaka, Malaysia

Artificial Immune Clonal Selection Based Algorithm in Academic Talent Selection

Research Article

Hamidah Jantan*, Siti 'Aisyah Sa'dan and Aisyah Mohd Faisal Baskaran

Faculty of Computer and Mathematical Sciences, University Teknologi MARA (Terengganu), Terengganu, Malaysia
***Corresponding author:** aisya.faisal@yahoo.com.my

Abstract. Artificial immune clonal selection based algorithms are centered on prominent immunological theory as computation mechanisms that emulate processes in biological immune system. Due to their abilities in achieving higher probability for global optimization, this technique has been used in many fields such as machine learning, computer security, pattern recognition, scheduling and etc. This article attempts to study the potential of clonal selection based algorithms for talent selection in higher learning institution. This study consists of three phases; data preparation for pre-processing academic talent data-sets; model construction using four selected artificial immune clonal selection based algorithms; and model analysis to measure the accuracy of the model proposed by the algorithms based on 10-fold cross validation. There were several experiments carried out using different set of training and testing data-sets; the significance test was also conducted in order to support the findings. As a result, *Clonal Selection Classifier* (CSCA) and *CLONal Selection ALGorithm* (CLONALG) algorithms proposed slightly high accuracy and would be considered as the potential classifiers for talent selection. For future work, the proposed model can be strengthened by the comparative study with other heuristic based searching algorithms in evolutionary computation paradigms such as genetic algorithm, ant colony optimization, swam algorithms and etc.

Keywords. Artificial Immune System; Clonal Selection Algorithms; Talent Selection

MSC. 03-XX

Received: April 5, 2016 **Accepted:** August 14, 2016

1. Introduction

Clonal selection based algorithm is known as the more flexible alternative to genetic and evolutionary algorithm in solving optimization problems. This algorithm uses colony search mechanism in nature and it can be converged to the global optimization with higher probability and speed [1–4]. Artificial immune clonal selection based algorithms are focused on prominent immunological theory as computation mechanisms that emulate processes in biological immune system which evolves candidate solutions by selecting, cloning and mutation procedures [5]. Clonal selection principle is the whole process of antigen recognition, cell proliferation and differentiation into memory cell. The main purpose of the algorithm is to develop a memory set of antibodies that represents a candidate's solution to a problem [6]. There are several artificial immune algorithms which have been developed imitating the clonal selection theory such as *CLONALG* (CLONal selection ALGorithm), *ICS* (Immunity Clonal Strategy algorithm, *Clonal Selection Classifier Algorithm* (CCSA), *Artificial Immune Recognition System* (AIRS), etc. [7, 8]. Besides that, this algorithm is in its canonical form and its various versions are used to solve different types of problems and are reported to perform better as compared to other heuristic methods such as genetic algorithm, neural network, ant colony optimization, etc. Due to the uniqueness ability in self-recognition, this algorithm has been mainly used by classifying self or non-self cells as detectors in function optimization, pattern recognition, design, scheduling, industrial engineering, fault diagnosis, machine learning, virus detection, etc. [8, 9]. In machine learning, clonal selection theory is used for learning and optimization to enhance the pattern matching process by classifying and clustering processes [10].

In any organization, the processes of managing people becoming an increasingly crucial method to approaching *Human Resource* (HR) management functions, which is known as talent management. The process of talent selection in HR field involves a lot of human decisions which are very subjective, uncertain and difficult especially in managing employee's advancement through promotion process. The talent selection process is a way for an organization to improve their employee's career path development which depends on several approaches and competency criteria [11]. Higher learning institutions is the places where knowledge is produced reproduced, preserved, systematized, organized and transmitted through academic activities conducted by the academia. Academia is known as an individual who has the competency to make a significant difference to current and future performance of higher learning institutions. The academic achievement is measured based on academic competency criteria related to his/her ability in academic activities such as teaching and supervision, research and publication, contribution to university and society, etc. This research attempts to study the potential of artificial immune clonal selection based algorithm as a mechanism for talent selection in *human resource* (HR) planning by classifying talent competencies.

This paper is discussed as follows: the second section explains the related work on the immune system *clonal selection algorith*m (CSA), clonal selection algorithm techniques and

application and academic talent selection in higher learning institution. The third section discusses the experimental phase conducted in this study. Then, the forth section explains the results and discussion. Finally, this paper ends with the conclusion remarks and future enhancements.

2. Related Works

2.1 Immune System Clonal Selection Algorithm

Clonal Selection Algorithm (CSA) using immunological principles used in *Artificial Immune System* (AIS) to describe the basic response of the adaptive immune system to an antigenic stimulus. The clonal selection theory is inspired by the immunology immune elements such as maintenance of specific memory set, selection and cloning of the most stimulated antibodies, death of non-stimulated antibodies, affinity maturation (mutation), reselection of the clones proportional with antigen, and diversity of the generation and maintenance [1, 12, 13]. The theory depends on the idea that only cells that are capable of recognizing an antigen will proliferate. The main purpose of this algorithm is to develop a memory set of antibodies that represents a solution to a problem.

2.2 CSA Techniques and Applications

CSA method is based on the biological feature in the immune system in the man's body [14]. CSA has two main computational mechanisms: selection and mutation. In the algorithm proposed by de Castro and Von Zuben, these two mechanisms were fulfilled by taking into account the immune properties: the proliferation and mutation rate is proportional to the antigenic affinity. The higher the antigenic affinity, the higher the number of clones generated for each antibody. In the hyper-mutation operation, the cloned population is subject to an affinity mutation process inversely proportional to the antigenic affinity. The receptor editing includes two steps. In the first step, a given number of new antibodies are generated randomly. In the second step, the generated antibodies are used to refresh the whole population by replacing those antibodies with the lowest antigenic affinity [12]. CSA uses three immune operators, i.e., cloning, hyper-mutation, and receptor editing, to refresh the composition of populations. The cloning operator explores the neighborhood of each point of the search space. This study employed four commonly used artificial immune classification algorithms: CLONALG, CSCA, AIRS1 and IMMUNOS for machine learning.

- **CLONALG:** Castro and Zuben proposed a clonal selection algorithm named CLONALG inspired by clonal selection theory for learning and optimization [15]. CLONALG generates a population of N antibodies, each specifying a random solution for the optimization process. Some of the best existing antibodies are selected, cloned and mutated in order to construct a new candidate population. New antibodies are then evaluated and certain

percentage of the best antibodies is added to the original population. Finally, a percentage of worst antibodies of previous generation is replaced with new randomly created ones [7]. The general steps of CLONALG were initially proposed to solve pattern recognition problems [8].

- **CSCA:** The CSCA was proposed by Brownlee in function optimization procedure that maximizes the number of patterns correctly classified and minimizes the number of patterns incorrectly classified [13]. There are several steps involved in CSCA technique such as antibody pool initialization, selection and pruning, cloning and mutation, clones and new antibody insertion, final pruning, and classification. CSCA is trained for several generations, and during each generation the entire set of antibodies is exposed to all antigens by classification process.
- **AIRS:** AIRS is a cluster-based approach to classification from data mining point of view. The AIRS algorithm has been shown to be a successful classification algorithm for a broad range of machine learning problems [10]. Initially, it learns the input space by mapping a cluster centers to it and then uses k-nearest neighbor on the cluster centers for classification. There are several processes involved in the algorithm such as memory cells stimulated by antigen proliferate and mutate. These stimulated cells and their progeny compete under selective pressure for continued stimulation, resulting in only the fittest being aggregated into the repertoire of "memory cell' that is used for classification. The attractive point of AIRS is its supervised procedure for discovering both the optimal number and position of the cluster centers [16].
- **IMMUNOS:** The first Immunos algorithm, known as Immunos-81 was proposed by Carter that uses the concept of artificial immune systems for classification problems. Immunos-81 algorithms are divided into two different configurations, which are Immunos-1 and Immunos-2, and an extension of over the algorithm is referred as Immunos-99 [17]. Immunos-1 algorithm uses no data reduction where the prepared clone population is kept to classify new data instances. In the training phase, a T-cell is prepared to represent the antigen-type and each group prepares a clone for the group allocated to the current T-cell. The creation of the B-cell population (clone) is a critical aspect of the algorithm as the clone is the basic recognition unit of the system. For each classification label, a B-cell population is created and for each antigen in the tanning set, a single B-cell is created. In the classification phase, for each B-cell clone population, an avidity value is determined. Based on the avidity value, a data instance is assigned to a class of clone with the highest avidity value handled by Euclidean distance measurement [3]. The Immnous-1 algorithm represents an instance-based classifier with some similarities to k-nearest neighbor classifiers.

CSA technique is actually a global optimization method inspired by biological clone selection principle to solve real world problems. Due to its ability in recognizing patterns

by proposing candidate's solutions for problem solving, it has been successfully applied into several challenging domains, such as multimodal optimization, machine learning and pattern recognition [14]. Table 1 shows several examples on general purpose of artificial immune clonal based selection algorithms to solve problem in various areas.

Table 1. Clonal Selection Algorithm Applications

Area	**Type of Application**
Pattern Recognition	Generalized Pattern Recognition [18]
	Handwritten Indic Script [19]
	Image classification Shape Recognition [20]
Optimization	Quantum Computing in Estimation of Distribution [21]
	Combinatorial Optimization in Travelling Salesman problem [22]
	Power Generators Maintenance Scheduling [4]
	Multi-objective Optimization Problem Clustering [23]
Machine learning	Binary Character Recognition in Engineering Applications [12]
	Function Optimization in Baldwinian Learning Operator [24]

2.3 Academic Talent Selection

In higher learning institutions, academic activities such as teaching, supervision, research, publication, etc. are known as the main duties for academics towards academic talent development and enhancement. Nowadays, academic talent marketplace is highly competitive to determine current and future direction for any education-based institutions [25]. Therefore, this issue is associated with the process of recruitment and retirement for academic talent as organization's key of long-term successfulness and competitiveness. The selection of academic talent based on their competencies in order to determine the right academic recognition for them such as professorial, scholarly, fellowship, expertise and etc. There are several studies conducted on this issue on talent selection by applying soft computing and data mining techniques [26, 27]. However, heuristic based searching algorithm in evolutionary computation, such as *Genetic Algorithm* (GA), *Ant Colony Optimization* (ACO), *Artificial Immune System* (AIS) and many others, has not attracted researchers in this area.

3. Experiment Setup

The first phase in this study is data preparation in order to collect and prepare academic achievement datasets. There were five datasets used in this project that were collected from several higher learning institutions, which represent the academic talent. Table 2 shows the description of attributes and number of data for each dataset.

Table 2. Dataset and Attribute Description

	Attribute	**Description**
Dataset1 (296)	Personal background (4)	Academic performance from academic Curriculum Vitae
	Academic Criteria (13)	
Dataset2 (658)	Personal background (3)	Yearly performance appraisal marks (for six years)
	Performance appraisal criteria (30)	
Dataset3 (198)	Personal background (4)	Publication achievement from academic Curriculum Vitae
	Publication criteria (6)	
Dataset4 (150)	Personal background (1)	Yearly performance appraisal marks on quality of work (from two evaluators)
	Performance appraisal criteria (25)	
Dataset5 (80)	Personal background (1)	Academic performance from academic promotion evaluation
	Promotion evaluation criteria (7)	

In the experimental phase, the model construction and model analysis were conducted using four selected artificial immune clonal selection based algorithms, i.e., CLONALG, CSCA, AIRS1 and Immunos-1. These algorithms were selected based on their strength and ability to optimize the proposed solution inspired by immunology system. In the model analysis, WEKA was used as machine learning tools to measure the accuracy of the proposed model using 10-fold cross validation process. There were several experiments carried out by using different set of training and testing data-sets to evaluate the accuracy of the proposed model. Finally, the significant test using paired T-Test was also applied for all selected algorithms and datasets in order to justify the results from both point of views (algorithm and data-set).

4. Results and Discussion

In this study, several experiments based on different datasets were conducted using selected clonal selection based algorithms, i.e., CLONALG, CSCA, AIRS1 and Immunos-1 using machine learning approach. Table 3 shows the accuracy of model using different clonal selection based algorithm for all datasets. The accuracy for the selected algorithms proposed medium rate of accuracy, i.e., 55%-70% for the datasets. The results show that the CSCA and CLONALG

algorithms proposed higher accuracy as compared to AIRS1 and Immunuos-1 algorithms. These algorithms,as what have been mentioned before, have their strength in learning and optimization for classification purpose.

Table 3. Accuracy of Model for Selected Algorithms

Dataset	**CLONALG**	**CSCA**	**AIRS1**	**IMMUNOS-1**
DS1	59.64	58.24	48.22	57.26
DS2	77.04	76.59	75.97	49.80
DS3	45.18	46.08	43.51	34.15
DS4	69.82	74.22	68.49	59.63
DS5	91.14	91.59	86.19	79.65
Average	68.56	69.34	64.48	56.10

AIRS1 and Immunos-2 algorithms are known as clustering based in machine learning as discussed before. Table 4 shows the accuracy of model for all datasets for different algorithms.

Table 4. Accuracy of Model for Selected Datasets

Dataset	**DS1**	**DS2**	**DS3**	**DS4**	**DS5**
CLONALG	59.64	77.04	45.18	69.82	91.14
CSCA	58.24	76.59	46.08	74.22	91.59
AIRS1	48.22	75.97	43.51	68.49	86.19
IMMUNOS-1	57.26	49.80	34.15	59.63	79.65
Average	55.84	69.85	42.23	68.04	87.14

The datasets have different attributes and number of data in representing academic talent criteria. The result shows that the dataset DS5 consists of 80 data with 8 attributes proposed high accuracy as compared to the others. This result indicates that the smaller the size of dataset and the number of attributes, the higher the accuracy of model will be produced. On the other hand, in order to propose the potential classifier for academic talent selection, the statistical test was conducted using t-test for significance analysis. Table 5 shows a positive mean difference in accuracy which shows that the CSCA algorithm has the highest value of positive mean which is significantly better than the other algorithms. For the accuracy criterion, the CSCA algorithm is significantly better than AIRS1 and Immunos-1 algorithms, with p-value < 0.05.

Table 5. Paired T-test on Accuracy of Model for Selected Algorithms

Paired Sample	Mean	SD	df	p-value
CSCA-CLONALG	0.780	2.207	4	.474
CSCA-AIRS 1	4.868	3.567	4	*.038
CSCA-Immunos 1	13.246	9.207	4	*.032

Besides that, for dataset distribution of attributes and number of data point of view, Table 6 shows a positive mean difference in accuracy which shows that the dataset DS5 has the highest value of positive mean which is significantly better than the other datasets (DS1, DS2, DS3 and DS4) with p-value < 0.05.

Table 6. Paired T-test on Accuracy of Model for Datasets

Paired Sample	Mean	SD.	df	p-value
DS5-DS1	31.302	6.535	3	*.002
DS5-DS2	17.292	8.624	3	*.028
DS5-DS3	44.912	1.503	3	*.000
DS5-DS4	19.102	1.891	3	*.000

5. Conclusion

In this study, clonal selection based algorithms were selected as the potential algorithm for academic talent selection in higher learning institution. As a result, CSCA and CLONALG algorithms would be considered as the potential technique due to the higher accuracy of model produced in model analysis phase. Besides that, AIRS1 and Immunos-1 algorithms that are commonly used for clustering would not be suitable for this type of datasets. From the dataset point of view, the attributes and number of data involved in model construction significantly affect the accuracy of the proposed model. Furthermore, the accuracy of model in this study can be strengthened by a comparative study with other heuristic based searching algorithm in evolutionary computation for classification such as genetic algorithm, ant colony optimization etc. As a conclusion, the ability of bio inspired algorithms for classification is continuously change and obtain new understanding in HR fields and become the major contribution to HR and machine learning technique.

Acknowledgement

This work has been supported by Ministry of Education (MOE) to Universiti Tekonologi MARA (UiTM) for Research Acculturation Grant Scheme (RAGS) (600-RMI/RAGS/5/2/(205/2013).

Competing Interests

The authors declare that they have no competing interests.

Authors' Contributions

All the authors contributed significantly in writing this article. The authors read and approved the final manuscript.

References

[1] K.A.A. Sheshtawi, H.M.A. Kader and N.A. Ismail, Artificial immune clonal selection classification algorithms for classifying malware and benign processes using API call sequences, *International Journal of Computer Science and Network Security* **10** (April 2010), 31–39.

[2] E.D. Ülker and S. Ülker, Comparison study for clonal selection algorithm and genetic algorithm, *Computer Science & Information Technology (IJCSIT)* **4** (August 2012), 107–118.

[3] A. Onnan, Artificial immune systems based web page classification, in *4th Computer Science Online Conference 2015* (*CSOC 2015*), On line Conference, 2015, pp. 189–199.

[4] M.Y. El-Sharkh, Clonal selection algorithm for power generators maintenance scheduling, *International Journal of Electrical Power & Energy Systems* **57** (2014), 73–78.

[5] W. Zhang, G.G. Yen and Z. He, Constrained optimization via artificial immune system, *IEEE Transactions on Cybernetics* **44** (February 2014), 185–198 .

[6] J. Brownlee, *Clonal Selection Theory & Clonalg the Clonal Selection Classification Algorithm (CSCA)*, Master of Information Technology, Faculty of Information & Communication Technologies (ICT), Swinburne University of Technology (2005).

[7] J.R. Al-Enezi, M.F. Abbod and S. Alsharhan, Artificial immune systems-models, algorithms and applications, *International Journal of Research and Reviews in Applied Sciences (IJRRAS)* **3** (2010), 118–131.

[8] B.H. Ulutus and S. Kulturel-Konak, A review of clonal selection algorithm and its applications, *Artificial Intelligence Review* **36** (2011), 117–138.

[9] A. Pachauri and Gursaran, Use of Clonal selection algorithm as software test data generation technique, in *2nd International Conference on Advanced Computing & Communication Technologies*, Rohtak, Haryana (2012), pp. 1–5.

[10] C. McEwan and E. Hart, On AIRS and clonal selection for machine learning, *Theoretical Computer Science* **412** (2011), 502–516.

[11] H. Jantan, A.R. Hamdan and Z.A. Othman, *Data Mining Classification Technique for Human Talent Forecasting* (2011 Edition), 21 January, 2011, pp. 1173–1178.

[12] L.N.D. Castro and F.J.V. Zuben, The clonal selection algorithm with engineering applications, in *Proceedings of GECCO*, Las Vegas, USA (2000), pp. 36–37.

[13] J. Brownlee, *Clonal Selection Theory & Clonalg The Clonal Selection Classification Algorithm (Csca)*, Master of Information Technology, Faculty of Information & Communication Technologies (ICT), Swinburne University of Technology, Swinburne University of Technology (SUT), 2005.

[14] R. Liu, P. Zhang and L. Jiao, Clonal Selection Classification Algorithm for High-Dimensional Data, in *LSMS / ICSEE 2010, Part II, CCIS 9* (editors K. Li et al.), pp. 89–95 (2010).

[15] L.N.D. Castro and J.I. Timmis, Artificial immune systems as a novel soft computing paradigm, *Soft Computing* **7** (2003), 526–544.

[16] A. Watkins, J. Timmis and L. Boggess, Artificial Immune Recognition System (AIRS): an immune-inspired supervised learning algorithm, *Genetic Programming and Evolvable Machines* **5** (2004), 291–317.

[17] J. Brownlee, *IMMUNOS-81 The Musunderstood Artificial Immune System, Centre for Intelligent Systems and Complex Processes (CISCP)*, Faculty of Information & Communication Technologies (ICT), Swinburne University of Technology (SUT) 2005.

[18] T. Huntsberger, Clonal selection based artificial immune system for generalized pattern recognition, in *2011 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Anchorage, AK 2011, pp. 3090–3095.

[19] U. Garain, M.P. Chakraborty and D. Dasgupta, Recognition of handwritten indic script using clonal selection algorithm, *ICARIS*, pp. 256-266, 2006.

[20] N. Isa, N.M. Sabri, K.S. Jazahanim and N.K. Taylor, Application of the Clonal Selection Algorithm in Artificial Immune Systems for Shape Recognition, 2010.

[21] O.S. Soliman and A. Rassem, A bio inspired clonal algorithm with estimationof distribution algorithm for global optimization, in *The 8th International Conference on INFOrmatics and Systems (INFOS2012) Bio-inspired Optimization Algorithms and Their Applications Track*, 2012.

[22] Y. Miao, Y. Yin and Y. Wang, Applications of clonal selection algorithm based on tabu criteria in combinatorial optimization, pp. 305-310, 2013.

[23] W. Ma, L. Jiao and M. Gong, Immunodominance and clonal selection inspired multiobjective clustering, *Progress in Natural Science* **19** (10 June 2009), 751–758.

[24] M. Gong, L. Jiao and L. Zhang, Baldwinian learning in clonal selection algorithm for optimization, *Information Sciences* **180** (2010), 1218–1236.

[25] P. Verhaegen, Academic talent: quo vadis? recruitment and retention of faculty in european business schools, *Journal of Management Development* **21** (2006), 807–818.

[26] H. Jantan, A.R. Hamdan and Z.A. Othman, Knowledge discovery techniques for talent forecasting in human resource application, in *World Academy of Science, Engineering And Technology*, Penang, Malaysia, 2009,

[27] J.-R. Chang, C.-H. Cheng and L.-S. Chen, A fuzzy-based military officer performance appraisal system, *Applied Soft Computing* **7** (2007), 936–945.