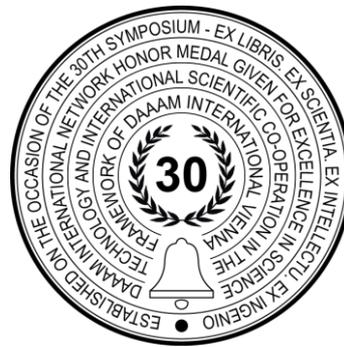


APPLICATION OF GENETIC ALGORITHMS IN HIGHER EDUCATION AREA

Danijel Kučak, Vedran Juričić & Goran Đambić



This Publication has to be referred as Kucak, D[anijel]; Juricic, V[edran] & Dambic, G[oran] (2019). Application of Genetic Algorithms in Higher Education Area, Proceedings of the 30th DAAAM International Symposium, pp.0343-0347, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-3-902734-22-8, ISSN 1726-9679, Vienna, Austria
DOI: 10.2507/30th.daaam.proceedings.045

Abstract

In computer science, a genetic algorithm (GA) is a computational method for solving optimization problems. Genetic algorithms are inspired by the process of natural selection that belongs to the larger class of evolutionary algorithms (EA). Genetic algorithms are commonly used to generate high-quality solutions to optimization and search problems by relying on bio-inspired operators such as mutation, crossover and selection. The application of genetic algorithms in higher education area is the focus of our study. The aim of this paper is to evaluate the possibilities of applying genetic algorithms in the higher education area. This paper identifies and analyses suitable literature, research papers and articles in order to determine their categorization in the field of higher education, to determine the current trends of using genetic algorithms in education, and to determine its current and future applications.

Keywords: genetic algorithm; higher education; resources allocation; student performance

1. Introduction

Genetic algorithm is a heuristic method of optimization that mimics the natural evolutionary process. [1]. There are lot of areas where generic algorithms could be applied, like traveling salesperson problem, decoding secret messaging, robot trajectory planning, etc. [2]

This study is focused on application of genetic algorithms in higher education. Education is changing on daily basis. There are not just students in classroom anymore, looking at the notebook, while a teacher lectures. Today's classrooms use digital resources [3]

In higher education, for example, genetic algorithms could be applied to support academic stuff creating timetable, analyse student performance, forecasting student enrolment, etc. The goal of our research was to categorize studies about the usage of genetic algorithms in higher education area. Before research has been started, we choose the research questions:

Research question 1: How to categorize research papers in the area of application of genetic algorithms in higher education?

Research question 2: What are trends in application of genetic algorithms in higher education area?

A comprehensive literature study has been made to get answers on these questions. We used the Systematic Literature Review (SLR) approach [5] to collect primary studies regarding this research scope. The main aim was to classify studies and the selected studies were then analysed and categorized using the content analysis method. As result of the classification, three main categories were defined:

- Improving student retention and student yield maximization
- Analysis of student performance
- Support to the academic staff

The structure of this study is as follows: In the second section, we reviewed genetic algorithms and its application in higher education area. In the third section, the research approach used in this study was introduced. After that, analysis of the results and categories of studies was introduced. Fifth section includes a conclusion and suggestions for future research on the topic.

2. Genetic algorithms and its application in higher education area

John H. Holland proposed genetic algorithms in the early seventies of last century. For more than four decades, they have proven to be very powerful and at the same time a general tool for solving a whole range of problems from engineering practice. This can be explained by their simplicity; both the very ideas on which they are based, and their applications; and the contribution of a number of scientists and engineers to their adaptation to a large number of problems and increasing efficiency [4]. Goldberg was probably the greatest evangelist of the GA with his book [6].

The analogy of evolution as a natural process and genetic algorithm as optimization methods is manifested in the selection process and genetic operators [1]. The mechanism of choice over a species of living beings in the evolutionary process makes the environment and conditions in nature.

In genetic algorithms, the key to selection is the “fitness function” (sometimes called “cost function”), which represents the problem being solved. As the environment in nature is key to selection over a species of living beings, so the fitness function is the key to selection of the population of solutions in the genetic algorithm. In the nature, an individual that best suits the conditions and environment in which he or she lives, the highest likelihood of survival and mating, and thus the transfer of his genetic material to his offspring.

Similarly, in genetic algorithm, selection of good individuals is transferred to the next population, and manipulation of genetic material creates new individuals. Such a cycle of selection, reproduction, and manipulation of genetic material of the individual is repeated until the condition for stopping the evolution process is met. The ultimate result is the population of individuals (potential solutions). The best one in the last iteration is the optimization solution.

The typical structure of sequential genetic algorithm, written as pseudocode (based on C programming language syntax):

```
init (P)
evaluation (P)
while not stop_criteria do
    P1 <- selection(P)
    crossover(P1)
    mutation(P1)
    P <- evaluation(P1)
end while
```

Each of steps could be implemented in different ways. In example, stop condition could be number of iterations executed, running time elapsed, etc. Authors in [2] extracted some advantages of generic algorithms:

Fig. 1. Genetic algorithm pseudocode

- Optimizes with continuous or discrete variables
- Deals with a large number of variables
- Is well suited for parallel computers
- Optimizes variables with extremely complex cost surfaces (they can jump out of a local minimum)
- Provides a list of optimum variables, not just a single solution
- May encode the variables so that the optimization is done with the encoded variables
- Works with numerically generated data, experimental data, or analytical functions.

3. Research

The above-mentioned research questions were extracted when we studied the aims of this research. To answer these research questions, we performed a literature study. We decided to use the SLR method to collect the relevant primary studies and followed the guidelines given by [5].

For the SLR, we decided to do an electronic search. The databases used were IEEE Xplore Digital Library (IEEE), Scopus database, Web of Science database (WoS), ScienceDirect, and others (Google Scholar and similar). The search term was (“Genetic algorithm” AND “Higher Education”) OR (“Genetic algorithm” AND “Student”). The targeted amount of related studies was between fifty and one hundred, because this amount would give us enough information for categorization and research trends.

In the next phase, we created three categories and classified the papers into those. The reason for this classification was that most of the papers published were relatively distinctive in terms of the research objective, methodology, and application. To be simplest as possible, we created three categories of studies without ignoring the variations of themes. This way, we examined the research papers that fell under the same. The final number of papers collected was 69. Some studies were included in more than just one category.

The research categories and the total number of selected studies are as follows:

- (A) Improving student retention and student yield maximization (9 studies)
- (B) Analysis of student data (42 papers)
- (C) Support to the academic staff (25 papers)

Letters A, B and C are just used as tags and serve as a mark for further referencing in the rest of this study.

4. Results

This section describes the main characteristics of each category. We present all three main categories in subsections, supported with relevant papers.

4.1. *Improving student retention and student yield maximization*

By identifying “at risk” students early, universities and colleges can detect and contact those students and help them to be more successful. Student retention is an essential part of many enrolment systems. It affects almost all segments of university metrics: reputation, financials, ranking. Specially, student retention has become one of the most important things for managers in higher education institutions. There are few studies, which developed models to predict and to introduce the reasons behind student’s number decreasing [7] [8] [9] [10].

4.2. *Analysis of student data*

Higher education institutions have to provide quality education to its students. One of the most popular way today of achieving highest level of quality in higher education system is by discovering knowledge for prediction regarding enrolment of students in a particular course, detection of fraud used in exams, detection of abnormal values in the results of the student deliverables, prediction about student performance, and so on. In general, higher education institutions collect large amounts of student data. The knowledge is hidden among the educational data set and genetic algorithm could be used to extract that information [11]. There are lot of papers which covers student data analysis in last years: [12] [14] [15].

4.3. *Support to the academic staff*

There are plenty of challenges in context of higher education where authors proposed models based on genetic algorithms. Creating resource allocation systems like university timetables, building recommending systems for students and automatized creation of assessment tests are some examples of solution based on genetic algorithms.

Higher education institutions are facing challenge when it comes to resources and events scheduling. Timetabling is such a kind of problem in which events (classes, exams, courses, students, professors, etc.) have to be allocated into a number of timeslots such that conflicts in using a given set of resources (classrooms, etc.) are avoided. Genetic algorithms are quite often used for creating automatized solutions to solve timetable creation challenge. Some of different models based on genetic algorithms are proposed in following papers: [17], [18], [19], [24] [13].

Automatized creation of assessment tests is another approach in higher education institutions having large number of students. Some studies propose genetic algorithm approach to that challenge: [20] [21].

Recommending systems become extremely interesting lately. Recommending elective courses to students or recommending learning paths to students are examples where model based on genetic algorithms are proposed lately [16] [22] [23].

5. Conclusion

The aim of this study was to evaluate the current state of the art in the application of genetic algorithms in education area. The amount of studies (papers and articles) was large, so only some of studies, which we found as good representatives, were mentioned in results this study. This study shows that there are significant different ways to benefit from genetic algorithms application in education area.

As we stated in introduction section, one of our goals was try to classify studies in the field of genetic algorithm application in education area. Based on our survey, the papers reviewed under category marked as A research a ways how genetic algorithms can improve students retention rate. Student retention is an essential part of many enrolment management systems. It affects university rankings, school reputation, and financial wellbeing. Student retention has become one of the most important priorities for decision makers in higher education institutions, so there are lot of studies in that category.

Reviewing studies under category marked as B, showed us how major benefit of genetic algorithms (regarding number of studies in scientific databases) is its ability to analyse large portion of student data. According to our survey, this is most interesting area of genetic algorithms application in higher education area to researchers. There are lot of studies in recent years in that category, and lot of models based on genetic algorithms were provided to analyse student data. We would say that this category is definitely the trend.

Reviewing studies under category marked as C, showed how genetic algorithms can support faculty staff in many ways. Creating timetable, build recommendation systems and automatized assessment creation are some examples. This is also quite interesting area, with lot of papers created lately.

As we have found earlier, research has been made over several relevant databases, but of course, not all were involved, so this can be considered as limitation of study. In addition, there is a possibility that some of the relevant studies may be skipped by chance.

In the future, we plan to implement timetable creating model based on genetic algorithms, for University College Algebra who has lot of studies, with some overlapping courses between different studies, increasing number of students and currently limited number of resources (classrooms and professors), so creating timetable is pretty challenging task for faculty staff. As none of the currently existing found timetable solutions tested in Algebra has not yielded satisfactory results, we hope that new model we plan to create, will.

6. References

- [1] Davis, L. (1991). Handbook of Genetic Algorithm. Van Nostrand Reinhold
- [2] Haupt, R.L. & Haupt S.E. (2004). Practical genetic algorithms. Wiley Intescience.
- [3] Kučak, D.; Juričić V. & Đambić, G. (2018). Machine learning in education – a survey of current research trends. doi:10.2507/29th.daaam.proceedings.059
- [4] Holland, B.A. & Charters, S. (1973) Genetic Algorithms and the Optimal Allocation of Trials, SIAM Journal on Computing
- [5] Kitchenham, B.A. & Charters, S. (2007) Guidelines for performing Systematic Literature Reviews in Software Engineering. Version 2.3, EBSE Technical Report EBSE-2007-01, Keele University, Keele, Staffordshire, UK
- [6] Goldberg, D.E. (1989) Genetic Algorithms in Search, Optimization, and Machine Learning. Reading, MA, Addison Wesley.
- [7] Yukselturk, E.; Ozekes, S. & Türell, Y.K. (2014). Predicting Dropout Student: An Application of Data Mining Methods in an Online Education Program. Computers in Human Behavior
- [8] Xing, W.; Guo, R.; Petakovic, E. & Goggins, S. (2015). Participation-based student final performance prediction model through interpretable Genetic Programming: Integrating learning analytics, educational data mining and theory. doi:10.1016/j.chb.2014.09.03
- [9] Sarafraz, Z.; Sarafraz, H.; Sayeh, M. & Nicklow, J. (2015). Student Yield Maximization Using Genetic Algorithm on a Predictive Enrollment Neural Network Model. doi:10.1016/j.chb.2014.09.034
- [10] Zahedifard, M.; Attarzadeh, I.; Pazhokhzadeh H. & Malekzadeh, J. (2015). Prediction of students' performance in high school by data mining classification techniques, International Academic Journal of Science and Engineering
- [11] Baradway, MB.K. & Pal, S. (2011). Mining Educational Data to Analyse Students' Performance. International Journal of Advanced Computer Science and Applications
- [12] Punithadevi, C.; Christina, E. A. & Manju, S. (2017). Analyzing student's academic performance using internal factors—A survey. Innovations in Power and Advanced Computing Technologies (i-PACT) (pp. 1-6).
- [13] Yemi, A. & Oludele O., (2019) A. Performance Evaluation of Genetic Algorithm Tuned With Particle Swarm Optimization for Solving University Examination Timetabling Problem.
- [14] Hamsa, H.; Indiradevi, S. & Kizhakkethottam, J. J. (2016). Student academic performance prediction model using decision tree and fuzzy genetic algorithm. Procedia Technology, 25, 326-332.
- [15] Kalles, D. & Pierrakeas, C. (2006). Analyzing student performance in distance learning with genetic algorithms and decision trees. Applied Artificial Intelligence, 20(8), 655-674.
- [16] Duan, X. (2019). Automatic Generation and Evolution of Personalized Curriculum Based on Genetic Algorithm. International Journal of Emerging Technologies in Learning, 14(12).

- [17] Matias, J. B.; Fajardo, A. C. & Medina, R. P. (2018) A Hybrid Genetic Algorithm for Course Scheduling and Teaching Workload Management. In 2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM) (pp. 1-6).
- [18] Srisamutr, A.; Raruaysong, T. & Mettanant, V. (2018). A Course Planning Application for Undergraduate Students Using Genetic Algorithm. In 2018 Seventh ICT International Student Project Conference (ICT-ISPC) (pp. 1-5).
- [19] Dener, M. & Calp, M. H. (2019). Solving the exam scheduling problems in central exams with genetic algorithms. arXiv preprint arXiv:1902.01360.
- [20] Ciguene, R.; Joiron, C. & Dequen, G. (2019). Automatically generating assessment tests within higher education context thanks to genetic approach. In *Bioinspired Heuristics for Optimization* (pp. 269-282). Springer, Cham.
- [21] Goularte, F. B.; Nassar, S. M.; Fileto, R. & Saggion, H. (2019). A text summarization method based on fuzzy rules and applicable to automated assessment. *Expert Systems with Applications*, 115, 264-275.
- [22] Esteban, A.; Zafra, A. & Romero, C. (2018). A Hybrid Multi-Criteria Approach Using a Genetic Algorithm for Recommending Courses to University Students. *International Educational Data Mining Society*.
- [23] Dwivedi, P.; Kant, V. & Bharadwaj, K. K. (2018). Learning path recommendation based on modified variable length genetic algorithm. *Education and Information Technologies*, 23(2), 819-836.
- [24] Modibbo, U. M.; Umar, I.; Mijinyawa, M. & Hafisu, R. (2019) Genetic Algorithm for Solving University Timetabling Problem.