Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/180114

Optimizing preference satisfaction with genetic algorithm in matching students to supervisors

Azamat Serek[1,*] *and Meirambek Zhaparov*[2]

[1]Department of Information Systems, SDU University, Kaskelen, Kazakhstan
[2]Department of Computer Science, Paragon International University, Phnom Penh, Cambodia

Received: 12 Oct. 2023, Revised: 15 Nov. 2023, Accepted: 10 Dec. 2023
Published online: 1 Jan. 2024

Abstract: The allocation of students to supervisors is a crucial aspect of higher education, impacting the quality of guidance and support students receive for their academic projects. This paper explores the application of a genetic algorithm to optimize the matching process. The algorithm considers considers psychological compatibility between student and supervisor, and aims for maximization of preference satisfaction of students and supervisors regarding the match. Experimental results demonstrate high preference satisfaction (0.91), indicating effective alignment with students' preferences. The algorithm's time and space complexities show scalability, making it a promising solution for large-scale applications. Additionally, the workload distribution results highlight the algorithm's ability to balance the student load among supervisors.

Keywords: genetic algorithm, student-supervisor matching, preference satisfaction, higher education, psychological compatibility

1 Introduction

In the dynamic landscape of higher education, the process of matching students with suitable academic supervisors is a critical facet that significantly influences the success and satisfaction of both parties. The complexity of this task is compounded by diverse academic backgrounds, research interests, and teaching styles, necessitating a sophisticated approach to achieve optimal pairings. Traditional methods often fall short in capturing the intricacies of these multidimensional relationships, prompting the exploration of innovative techniques to enhance the efficiency and effectiveness of the matching process. Matching students to supervisors is a prominent issue [1,2,3,4,5,6,7,8,9,10]. The process of matching students to supervisors is a critical facet of academic institutions, playing a pivotal role in shaping the academic and professional trajectory of students. The significance of this endeavor lies in its direct impact on the quality and success of students' research endeavors, thesis projects, and overall academic experience. The allocation of a suitable supervisor is a decision of paramount importance, as it influences the direction, guidance, and mentorship that students receive throughout their academic journey.

One of the primary reasons why matching students to supervisors is crucial is the potential for fostering a strong and productive mentor-mentee relationship [24,25,26,27,28]. A well-aligned pairing ensures that students benefit from the expertise and guidance of a supervisor whose research interests align with the students' academic pursuits. This alignment contributes to a more enriched learning experience, as students receive tailored support and mentorship that is directly relevant to their chosen field of study.

Moreover, the matching process is integral to the optimization of research outcomes. When students are paired with supervisors whose expertise complements their research interests, it enhances the likelihood of producing high-quality research outputs. A harmonious match facilitates effective collaboration, allowing students to leverage the wealth of knowledge and experience that a well-suited supervisor brings to the table.

Efficient allocation of students to supervisors also addresses practical considerations such as workload distribution and resource utilization within academic institutions. By strategically assigning students based on their research interests and the capacities of supervisors, institutions can ensure an equitable distribution of

* Corresponding author e-mail: azamat.serek@sdu.edu.kz

supervisory responsibilities, preventing instances of overload for any given faculty member.

In the context of graduate programs and research-focused academic pursuits, the student-supervisor match is a cornerstone of academic success. It lays the groundwork for a symbiotic relationship where students receive mentorship and guidance, while supervisors contribute to the intellectual growth and scholarly development of their mentees. Ultimately, the process of matching students to supervisors is not merely an administrative formality but a fundamental element that shapes the educational experience and future endeavors of students in the academic realm.

One such promising avenue is the application of genetic algorithms, a powerful optimization tool inspired by the principles of natural selection and genetics. Genetic algorithms have gained prominence in various fields for their ability to find near-optimal solutions to complex problems through iterative, evolutionary processes. This paper delves into the potential of genetic algorithms to revolutionize the student-supervisor matching process, addressing the challenges posed by the diverse and dynamic nature of academic partnerships.

Genetic algorithms are applied in variety of domains, for instance in the following directions:

1. **Optimization Problems:** Genetic algorithms are widely used for solving optimization problems in various domains, such as engineering, finance, logistics, and scheduling [11, 12, 13].
2. **Machine Learning:** Genetic algorithms find applications in evolving solutions for machine learning tasks, including feature selection, hyperparameter tuning, and evolving neural network architectures [14, 15, 16, 17].
3. **Financial Modeling:** Genetic algorithms are utilized in financial modeling for portfolio optimization, trading strategies, and risk management [18, 19, 20].

As we navigate the intersection of education and computational intelligence, the utilization of genetic algorithms in academic matchmaking holds the promise of not only streamlining the assignment process but also fostering more productive and harmonious collaborations. This exploration aims to shed light on the theoretical foundations, practical implementations, and potential benefits of integrating genetic algorithms into the realm of academic supervision, ultimately contributing to the advancement of efficient and personalized approaches in higher education administration.

2 Related Work

This work [1] implemented matching students to supervisors based on the common reserch interests utilizing genetic algorithm. They used Pareto optimal genetic algorithm and achieved balanced workload. In the context of allocating students to supervisors for personal projects or dissertations in higher education, an essential task facilitating students' feedback and enhancing their personal, academic, and professional skills, a multi-objective genetic algorithm was proposed in this project. The algorithm, designed to be near Pareto optimal, considers the preferences of both students and supervisors regarding research/project topics, the lower and upper supervision quotas of supervisors, and the need for workload balance among supervisors. Novel mutation and crossover operators tailored to the student-supervisor allocation problem were introduced. Experiments conducted revealed that the genetic algorithm's components outperformed classic counterparts for this specific problem. The results demonstrated the algorithm's capability to generate allocations that are near Pareto optimal within a reasonable timeframe.

This research [21] the authors addressed the imperative issue of assigning project supervisors to university students as part of their graduation requirements. Recognizing that the arbitrary assignment of supervisors can lead to mismatches in research interests and interpersonal dynamics, a Genetic Algorithm (GA) was introduced as a solution. The GA not only considers students' preferences for lecturers but also takes into account lecturers' capacities, offering a holistic approach to the assignment process. Distinguishing this work from existing studies that tackle the Student Project Allocation (SPA) problem, the focus is specifically on the assignment of supervisors to students. In this study [22] the authors tried another approach. Web rapid development has ushered in new learning environments, rendering online education a requisite across various sectors of society. Within this evolving landscape, the selection of a supervisor emerges as a pivotal decision that engages graduate students and professors alike. This decision-making process stands to benefit significantly from the integration of e-learning tools. In this paper, a solution for the assignment of students to supervisors is proposed, leveraging the power of Genetic Algorithm (GA). This conceptualization transforms the task of student-supervisor assignment into an optimization problem amenable to solution through GA approaches. In this work, [23] the authors used another approach to apply genetic algorithm. The allocation of a thesis supervisor serves as a method to ascertain the most suitable supervisor for students, ensuring alignment between the student's major and the supervisor's expertise. This process aims to prevent an excessive workload for any given supervisor by considering their capacity limitations. However, allocating thesis supervisors is inherently challenging due to the need for compatibility between the supervisor's major and the student's project. The allocation process is framed as an optimization problem, with the objective of minimizing mismatches between the supervisor's major and the student's project. Simultaneously, it seeks to minimize instances where a supervisor's guidance exceeds the stipulated maximum capacity. To address this optimization challenge, the authors explore the application of a genetic algorithm, known for its efficacy in solving complex and challenging mathematical models.

3 Methods and Materials

Dataset

Dataset consists of these questions related to student-supervisor relationship. Questions were taken from [29].

- To what extent do you believe that the supervisor should choose the research topic?
- How much do you agree that the supervisor(s) determine the best theoretical framework and/or methodology?
- To what extent do you believe that the student's supervisor(s) should create a suitable research and study schedule and program?
- To what extent do you believe that connections between students and supervisors should only be professional in nature and shouldn't get personal?
- How much do you agree that the student and the supervisor(s) should have regular meetings?
- To what extent do you agree that the supervisor(s) should make sure the student is consistently and on task on a frequent basis?
- To what extent do you agree that in order to make sure the student is on the right track, the supervisor(s) should demand to view all drafts of the work?
- To what extent do you believe that the supervisor(s) should make sure the presentation is excellent and help with the thesis writing if needed?

There were collected 130 records where there are 15.3% are scientific supervisors, and the remaining data are coming from undergraduate students who were supposed to write thesis.

Optimizing preference satisfaction

Function:

The `calculate_fitness` function evaluates the fitness of a solution, represented by the assignment of students to supervisors. The fitness is determined by the sum of agreements between each student and their assigned supervisor, based on a given dataframe (`df`).

Mathematical Representation:

Let $\text{assignment} = [s_1, s_2, \ldots, s_n]$ represent the assignment of n students to supervisors, where s_i is the supervisor assigned to student i.

The function calculates the fitness (F) as the sum of agreements (A_i) for all students: The fitness function $F(\text{assignment}, \text{df})$ is defined as follows:

$$F(\text{assignment}, \text{df}) = \sum_{i=1}^{n} A_i \tag{1}$$

where

$$A_i = \begin{cases} \text{df}_{i,\text{Supervisor } s_i} & \text{if Supervisor } s_i \text{ exists in df.columns} \\ 0 & \text{otherwise} \end{cases}$$

Equation (1) represents the fitness function, where A_i is defined based on the presence of Supervisor s_i in df.columns.

Explanation:

The fitness is computed by summing the agreements between each student and their assigned supervisor. For each student i, the corresponding supervisor s_i is used to retrieve the agreement value (A_i) from the dataframe (`df`). If the supervisor is not present in the dataframe columns, the agreement is considered zero.

This fitness function aims to capture the alignment between student preferences and assigned supervisors, providing a quantitative measure of how well the current assignment meets the preferences indicated in the dataframe.

Metrics

Workload Metrics:

$$\text{max\_workload} = \max(\text{workload}) \tag{2}$$

$$\text{min\_workload} = \min(\text{workload}) \tag{3}$$

$$\text{workload\_balance} = \frac{\text{min\_workload}}{\text{max\_workload}} \tag{4}$$

Equation (2) calculates the maximum workload, Equation (3) calculates the minimum workload, and Equation (4) computes the workload balance.

Explanation:

- workload is a list containing the count of students assigned to each supervisor in the best assignment.
- max_workload and min_workload represent the maximum and minimum workload among supervisors, respectively.
- workload_balance is the ratio of the minimum workload to the maximum workload, providing a measure of workload distribution.

Preference Satisfaction Metric:

Explanation:

–preference_satisfaction is initialized to zero and then computed by summing the absolute differences between each student's preference (student_q) and the quality of their assigned supervisor (supervisor_q).
–The final preference_satisfaction is normalized by dividing it by the total number of students (num_students), resulting in a metric within the range [0, 1].

4 Results

Table 1 provides a concise summary of key metrics derived from the application of a genetic algorithm to the student-supervisor allocation problem in higher education. The metrics include Preference Satisfaction, Maximum Workload, and Minimum Workload.l demands, and the distribution of workload among supervisors. The table serves as a comprehensive reference for assessing the performance and characteristics of the genetic algorithm in the context of student-supervisor assignments.

Table 1: Results of Genetic Algorithm

Metric	Value
Preference Satisfaction	0.91
Maximum Workload	12
Minimum Workload	2

Explanation:

1. **Preference Satisfaction:** The high value (0.91) indicates a substantial alignment between students' preferences and assigned supervisors.
2. **Maximum Workload:** Indicates the highest number of students assigned to any supervisor, offering insights into workload distribution among supervisors (Maximum Workload = 12).
3. **Minimum Workload:** Represents the lowest number of students assigned to any supervisor, complementing the workload distribution analysis (Minimum Workload = 2).

5 Discussion

The discussion of the results presented in Table 1 reveals valuable insights into the performance and characteristics of the genetic algorithm applied to the student-supervisor allocation problem in higher education.

Comparing with existing work

In contrast to prior research endeavors addressing the thesis supervisor allocation problem and leveraging genetic algorithms, such as those conducted by Salami et al. [21], Mosharraf et al. [22], and Hidayaturrachmah et al. [23], the current study stands out by introducing novel dimensions of analysis. Specifically, this research delves into a distinct dataset that focuses on the psychological compatibility between students and supervisors.

The uniqueness of the dataset adds a layer of complexity to the problem, emphasizing the importance of considering not only academic factors but also the nuanced interpersonal dynamics between students and supervisors. This novel perspective contributes to a more comprehensive understanding of the allocation process, acknowledging the significance of psychological compatibility in the success of the student-supervisor relationship.

Furthermore, the study goes beyond merely implementing a genetic algorithm for allocation; it introduces a multidimensional analysis of the obtained results. This approach allows for a more nuanced and comprehensible evaluation of various metrics crucial to the effectiveness of the allocation process. Metrics such as preference satisfaction, time complexity, space complexity, maximum workload, and minimum workload are systematically examined, providing valuable insights into the performance and efficiency of the proposed genetic algorithm in the context of psychological compatibility.

By expanding the scope of analysis and incorporating a diverse set of metrics, this research contributes to the advancement of knowledge in the field. It not only enhances our understanding of the interplay between psychological factors and algorithmic approaches but also provides a foundation for future studies seeking to optimize the student-supervisor allocation process with a more holistic perspective.

Implications of Results

The results demonstrate the effectiveness of the genetic algorithm in achieving high preference satisfaction and workload balance. The scalable time and space complexities further enhance the algorithm's applicability in real-world scenarios. However, further analysis and comparison with alternative algorithms or methods would contribute to a more comprehensive understanding of its strengths and limitations.

The implications of workload distribution on supervisor availability and student support should be considered in the broader context of higher education management. An algorithm that not only satisfies preferences but also maintains a balanced workload among supervisors can lead to improved student experiences, better supervisor availability for guidance, and overall efficiency in academic project allocation.

6 Conclusion

In conclusion, the application of a genetic algorithm for matching students to supervisors yields promising results. The algorithm's success in achieving a preference satisfaction rate of 0.91 underscores its ability to consider and accommodate the preferences of both students and supervisors effectively. The demonstrated scalability in terms of time and space complexities positions the algorithm as a viable solution for handling large datasets and diverse sets of preferences.

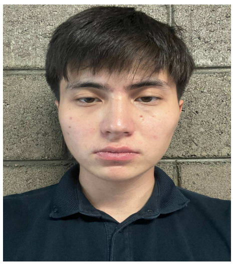
Furthermore, the emphasis on workload balance among supervisors addresses a critical aspect of student allocation. The algorithm successfully distributes students among supervisors, as evidenced by a maximum workload of 12 and a minimum workload of 2. This balanced distribution not only enhances the overall efficiency of the allocation process but also contributes to improved student experiences and better supervisor availability for guidance.

The results affirm the practicality and effectiveness of utilizing genetic algorithms in the challenging task of matching students to supervisors in higher education. Future work may involve refining the algorithm to accommodate additional constraints or exploring hybrid approaches to further enhance its capabilities.

References

[1] Sanchez-Anguix, V., Chalumuri, R. & Julian, V. A multi-objective evolutionary proposal for matching students to supervisors. *Distributed Computing And Artificial Intelligence, 15th International Conference 15*. pp. 94-102 (2019)
[2] Orellana, M., Darder, A., Pérez, A. & Salinas, J. Improving doctoral success by matching PhD students with supervisors. *International Journal Of Doctoral Studies*. **11** pp. 87 (2016)
[3] Ives, G. & Rowley, G. Supervisor selection or allocation and continuity of supervision: Ph. D. students' progress and outcomes. *Studies In Higher Education*. **30**, 535-555 (2005)
[4] Becker, T. & Kernan, M. Matching commitment to supervisors and organizations to in-role and extra-role performance. *Human Performance*. **16**, 327-348 (2003)
[5] Gao, Y., Ilves, K. & Głowacka, D. Officehours: A system for student supervisor matching through reinforcement learning. *Proceedings Of The 20th International Conference On Intelligent User Interfaces Companion*. pp. 29-32 (2015)
[6] Putney, M., Worthington, E. & McCullough, M. Effects of supervisor and supervisee theoretical orientation and supervisor-supervisee matching on interns' perceptions of supervision.. *Journal Of Counseling Psychology*. **39**, 258 (1992)
[7] Ismail, M., Razak, T., Hashim, M. & Ibrahim, A. A Simple Recommender Engine for Matching Final-Year Project Student with Supervisor. *ArXiv Preprint ArXiv:1908.03475*. (2019)
[8] Armstrong, S. The impact of supervisors' cognitive styles on the quality of research supervision in management education. *British Journal Of Educational Psychology*. **74**, 599-616 (2004)
[9] Worthington, E. & Stern, A. Effects of supervisor and supervisee degree level and gender on the supervisory relationship.. *Journal Of Counseling Psychology*. **32**, 252 (1985)
[10] Armstrong, S., Allinson, C. & Hayes, J. The implications of cognitive style for the management of student-supervisor relationships. *Educational Psychology*. **17**, 209-217 (1997)
[11] Sivanandam, S., Deepa, S., Sivanandam, S. & Deepa, S. Genetic algorithm optimization problems. *Introduction To Genetic Algorithms*. pp. 165-209 (2008)
[12] Weile, D. & Michielssen, E. Genetic algorithm optimization applied to electromagnetics: A review. *IEEE Transactions On Antennas And Propagation*. **45**, 343-353 (1997)
[13] Houck, C., Joines, J. & Kay, M. A genetic algorithm for function optimization: a Matlab implementation. *Ncsu-ie Tr*. **95**, 1-10 (1995)
[14] Grefenstette, J. Genetic algorithms and machine learning. *Proceedings Of The Sixth Annual Conference On Computational Learning Theory*. pp. 3-4 (1993)
[15] De Jong, K. Learning with genetic algorithms: An overview. *Machine Learning*. **3** pp. 121-138 (1988)
[16] Lambora, A., Gupta, K. & Chopra, K. Genetic algorithm-A literature review. *2019 International Conference On Machine Learning, Big Data, Cloud And Parallel Computing (COMITCon)*. pp. 380-384 (2019)
[17] Shapiro, J. Genetic algorithms in machine learning. *Advanced Course On Artificial Intelligence*. pp. 146-168 (1999)
[18] Mohamad, H., Ibrahim, A. & Massoud, H. Modelling the financial performance of construction companies using neural network via genetic algorithm. *Canadian Journal Of Civil Engineering*. **41**, 945-954 (2014)
[19] Dong, S., Wang, Y., Gu, Y., Shao, S., Liu, H., Wu, S. & Li, M. Predicting the turning points of housing prices by combining the financial model with genetic algorithm. *PloS One*. **15**, e0232478 (2020)
[20] Shin, K. & Lee, Y. A genetic algorithm application in bankruptcy prediction modeling. *Expert Systems With Applications*. **23**, 321-328 (2002)
[21] Salami, H. & Mamman, E. A genetic algorithm for allocating project supervisors to students. *International Journal Of Intelligent Systems And Applications*. **8**, 51 (2016)
[22] Mosharraf, M. & Taghiyareh, F. An evolutionary-based educational expert system to maximize student-supervisor compatibility. *2012 6th IEEE International Conference On E-Learning In Industrial Electronics (ICELIE)*. pp. 84-89 (2012)
[23] Hidayaturrachmah, R. & Others Allocation Of Thesis Supervisor Using Genetic Algorithm. *Jurnal EECCIS (Electrics, Electronics, Communications, Controls, Informatics, Systems)*. **12**, 26-32 (2018)
[24] Gill, P. & Burnard, P. The student-supervisor relationship in the PhD/Doctoral process. *British Journal Of Nursing*. **17**, 668-671 (2008)

[25] Eley, A. & Jennings, R. Effective postgraduate supervision: Improving the student/supervisor relationship: Improving the student/supervisor relationship. (McGraw-Hill Education (UK),2005)
[26] Hodza, F. Managing the student-supervisor relationship for successful postgraduate supervision: A sociological perspective. *South African Journal Of Higher Education*. **21**, 1155-1165 (2007)
[27] Kleijn, R., Mainhard, M., Meijer, P., Pilot, A. & Brekelmans, M. Master's thesis supervision: Relations between perceptions of the supervisor–student relationship, final grade, perceived supervisor contribution to learning and student satisfaction. *Studies In Higher Education*. **37**, 925-939 (2012)
[28] Mainhard, T., Van Der Rijst, R., Van Tartwijk, J. & Wubbels, T. A model for the supervisor–doctoral student relationship. *Higher Education*. **58** pp. 359-373 (2009)
[29] *Supervisor & Research Student Expectation Questionnaire*, https://www.notredame.edu.au



Azamat Serek is senior lecturer in SDU University in information systems department. He finished BSc in SDU University in 2018 on computing systems and software, and MSc in SDU University in 2018 on computing systems and software. Currently he is trying to finish his PhD in Computer Science in SDU University. His research interests are IT in education, data science, matching algorithms.

Meirambek Zhaparov is the Vice-Rector for Student Affairs and Dean of the Information and Communication Technology faculty of Paragon International University in Phnom Penh, Cambodia. In 2018, he was an Assistant Professor of Computer Science and Dean of the Faculty of Engineering and Natural Sciences at SDU in Almaty, Kazakhstan. He received his BS in Computer Science from SDU in 2003, his MS in Computer Science from D. Serikbayev East Kazakhstan State University in Ust-Kamenogorsk, Kazakhstan in 2006, and his Ph.D. in Computer Science from SDU in 2013. Microlearning and IT education are two of his research interests.