13th International Conference on Applied Informatics
Eszterházy Károly Catholic University
Eger, Hungary, February 19–21, 2026

Fair Team Formation for STEAM Education Using Exact and Heuristic Constraint-Based Solvers

Ali Adil Adil[a], **Gergely Kovásznai**[b], **Marcell Herbák**[b], **Mustafa Khalid Mezaal**[c]

[a]University of Debrecen
ali.adil@inf.unideb.hu

[b]Eszterházy Károly Catholic University
kovasznai.gergely@uni-eszterhazy.hu
herbakmarcell@gmail.com

[c]Imam Jaafar Al-Sadiq University
mustafa.khalid@ijsu.edu.iq

Abstract

STEAM education emphasizes interdisciplinary learning by integrating science, technology, engineering, arts, and mathematics. Collaborative, project-based activities are a core component of STEAM pedagogy, making fair and balanced team formation an important challenge. This paper studies the problem of automated team formation for STEAM education under hard constraints on skills, gender balance, and disciplinary diversity. We compare three solver-based approaches: an SMT-based solver using Z3, a constraint programming solver using CP-SAT, and a heuristic Genetic Algorithm.

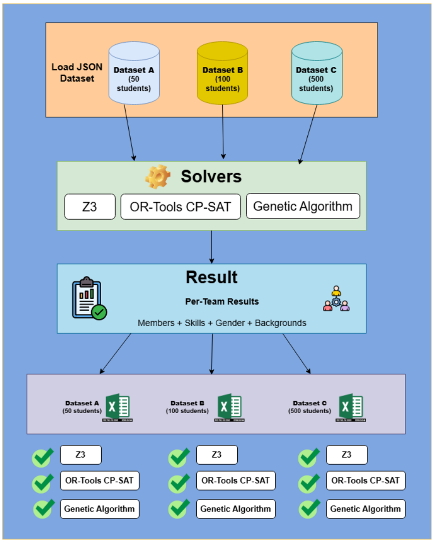
Experiments are conducted on a synthetic dataset [2], designed to reflect STEAM-oriented skill distributions. The dataset consists of student records and project requirements, including team size, minimum STEAM skill thresholds, and diversity constraints. Student attributes include gender, academic background (Engineering, Arts, or Science) and STEAM related skill ratings in Programming, Design, Math, Creativity, and Science. All skills are discretized on a scale from 1 (lowest) to 5 (highest) to ensure consistent processing during team formation.

The team formation model is defined using a fixed set of constraints derived from

STEAM education requirements. The specific parameters of the model, including team size, minimum skill thresholds, and diversity constraints, are summarized in Table 1a. The dataset is processed along the workflows in Figure 1b, where student data are passed to the solvers under the same constraint configuration. This formulation ensures that all solvers operate on an identical model, enabling a fair and consistent comparison of their ability to generate balanced STEAM teams.

Team size	5 members
Min. skills	Programming: 10, Design: 10, Math: 8, Creativity: 6, Science: 8
Min. diversity	**Gender**: - Female (F): min. 2 - Male (M): min. 2 **Backgrounds**: - Engineering - Arts - Science

(a) Model constraints.



(b) Three workflows.

Figure 1. Model constraints and workflows for team formation.

Preliminary experiments were conducted on three datasets containing 50, 100, and 500 students, respectively, to evaluate solver scalability and team quality. All solvers were executed under identical model settings with a global timeout of 100 seconds, ensuring a fair comparison. Across all datasets, CP-SAT [1] of Google's OR-Tools achieved the best overall balance between efficiency and scalability, consistently requiring the lowest total runtime, as being illistrated in Figure 2. Z3 exhibited fast reasoning but incurred increasing overhead as dataset size grew, leading to significantly higher runtime for larger instances. The Genetic Algorithm [3] consistently produced feasible solutions and scaled reliably, although it required more time than OR-Tools.

In addition to runtime, team quality was assessed using minimum and maximum total STEAM skill scores. As shown in Table 1, exact solvers produced more balanced teams with smaller Min–Max gaps, while the Genetic Algorithm showed larger variation between teams, particularly for larger datasets.

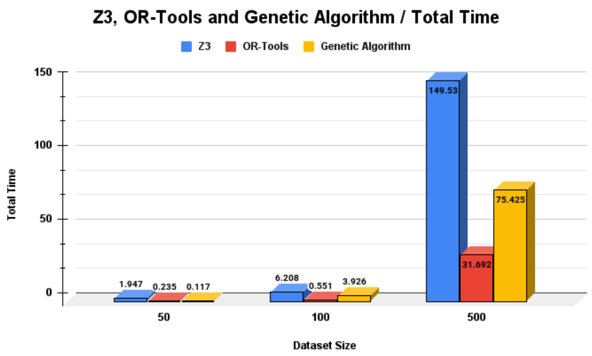


Figure 2. Runtime comparison of Z3, OR-Tools (CP-SAT), and Genetic Algorithm on synthetic datasets (50, 100, 500 students), including preprocessing overhead.

Table 1. Minimum and maximum total STEAM skill scores and fairness gap across datasets.

Dataset Size	Solver	Min STEAM	Max STEAM	Gap
50	Z3	78	83	5
50	OR-Tools	75	82	7
50	Genetic Algorithm	77	82	5
100	Z3	73	88	15
100	OR-Tools	69	91	22
100	Genetic Algorithm	72	86	14
500	Z3	74	85	11
500	OR-Tools	75	87	12
500	Genetic Algorithm	70	84	14

The results highlight a clear trade-off between scalability and fairness in STEAM-oriented team formation, and the future work will focus on hybrid solver strategies, the inclusion of optimization objectives beyond feasibility, and validation using real-world STEAM education datasets.

References

[1] B. Lennartson: *Optimization of Timed Petri Nets using CP-SAT*, IFAC-PapersOnLine 58.1 (2024), 17th IFAC Workshop on discrete Event Systems WODES 2024, pp. 90–95, ISSN: 2405-8963, DOI: https://doi.org/10.1016/j.ifacol.2024.07.016.

[2] Q. Liu, R. Shakya, J. Jovanovic, M. Khalil, J. de la Hoz-Ruiz: *Ensuring privacy through synthetic data generation in education*, British Journal of Educational Technology 56 (Feb. 2025), pp. 1053–1073, DOI: 10.1111/bjet.13576.

[3] J. Moreno, D. A. Ovalle, R. M. Vicari: *A genetic algorithm approach for group formation in collaborative learning considering multiple student characteristics*, Computers and Education 58.1 (2012), pp. 560–569, DOI: https://doi.org/10.1016/j.compedu.2011.09.011.