Forming Heterogeneous Groups for Intelligent Collaborative Learning Systems with Ant Colony Optimization

Sabine Graf
Vienna University of Technology
Austria
graf@wit.tuwien.ac.at

Rahel Bekele
Addis Ababa University
Ethiopia
rbekele@sisa.aau.edu.at

Motivation and Aim

- Collaborative learning is one of the many instructional approaches to enhance student performance
- Collaborative learning has many advantages
- Computer-based tools for collaborative learning focus mainly on collaborative interaction (sharing information & resources between students)
- Group formation process plays a critical role

  → heterogeneity

Aim:

Develop a tool that supports group formation by incorporating heterogeneity based on personality and performance attributes

→ Mathematical approach for the group formation problem

→ Optimization algorithm (Ant Colony Optimization)

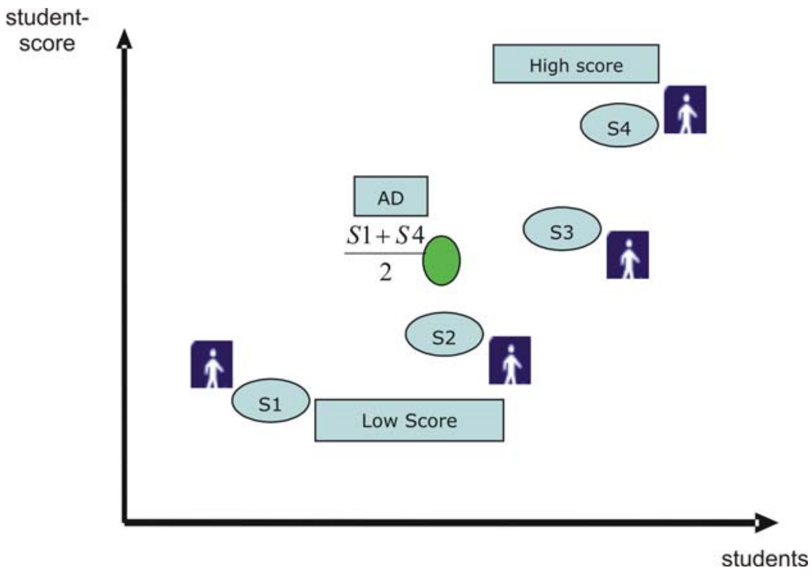
→ Experiments on developed tool

Mathematical Approach

- Personality and performance attributes:
  - Group work attitude
  - Interest for the subject
  - Achievement motivation
  - Self-confidence
  - Shyness
  - Level of performance in the subject
  - Fluency in the language of instruction
- Each attribute has three values
  (1= low, 2 = moderate, 3 = high)
- Vector space model for describing students' data
  e.g.: $S_1(3, 1, 2, 1, 3, 3, 2)$
- Student score: $\sum_{i=1}^{n} A_i\,(S_j)$
- Heterogeneity between two students: Euclidean Distance (ED)

Goodness of Heterogeneity (GH)

- Small, mixed-ability groups of four members:
  1 high achiever, 2 average achievers, and 1 low achiever (Slavin, 1987)



$$AD_i = \frac{\max scoreof(S_1, S_2, S_3, S_4) + \min scoreof(S_1, S_2, S_3, S_4)}{2}$$

$$GH_i = \frac{\max scoreof(S_1, S_2, S_3, S_4) - \min scoreof(S_1, S_2, S_3, S_4)}{1 + \sum_j \left| AD - scoreof(S_{j(i)}) \right|}$$

Forming Heterogeneous Groups

- Previous experiment:
  - Students were grouped randomly, on self-selection basis, or according to GH

  → Students who are grouped according to GH performed better
- Limitation of GH: based only on score values

  S_1 (3, 1, 2, 1, 3, 3, 2) → student score = 15

  S_2 (1, 3, 3, 2, 1, 2, 3) → student score = 15
- Extended approach
  - Groups should have high, average, and low achiever (GH)
  - Incorporate personality and performance attributes separately (Euclidean Distance)
  - Groups with similar degree of GH → coefficient of variation (CV) of GH values

  → Objective function:

$$F = w_{GH} \cdot GH + w_{CV} \cdot CV + w_{ED} \cdot ED \rightarrow \max$$

- Multi-agent meta-heuristic for solving NP-hard combinatorial optimization problems
- Advantages
  - Easy to apply to different optimization problems (only requirement: representation as graph)
  - Algorithm can be adapted to the problem rather than adapting the problem to the algorithm
  - Decentralization and indirect communication
- Ant Colony System
  - Developed by Dorigo and Gambardella (1997a)
  - Competitive with other optimization approaches such as neural networks and genetic algorithms (Dorigo and Gambardella, 1997b)

Ant Colony Optimization – Basic Concept

- Trail-laying trail-following behavior
  - Ants lay pheromone trails
  - Succeeding ants decide about the next node based on local and global information (random proportional transition rule)
  - The more pheromones on a path, the greater the probability that succeeding ants use this path, which lay again pheromones
- Pheromones evaporate over time

→ Global map of pheromone trails
(indicating the quality of the paths)

Applying ACS to the Group Formation Problem

- How to calculate local information?
  - Euclidean Distance (ED)
  - Goodness of heterogeneity (GH)
- How to calculate global information?
  - Based on the approach in ACS (pheromone update rules)
  - Updating is done between all edges in the group (amount of pheromones is for each of these edges equal)
- How to measure the quality of the solution?
  - 2-opt local search method is applied to each solution
  - Quality is measured according to the objective function

$$F = w_{GH} \cdot GH + w_{CV} \cdot CV + w_{ED} \cdot ED \rightarrow \max$$

Experiments

- 512 student data records
- 5 randomly chosen data sets of 100 students
- 20 runs per data set
- Each run is performed at least for 100 iterations and stops after the solution does not changed over the last 2/3 iterations
- Result:

Dataset	No. of students	Average GH	Average CV	Average ED	Average Fitness	SD Fitness	CV Fitness
A	100	129.81286	39.22323	363.93597	52.14131	0.03320	0.06367
B	100	117.20000	35.18174	377.41486	51.55805	0.02935	0.05693
C	100	114.23423	41.90564	374.14736	49.42179	0.03290	0.06656
D	100	132.17583	31.34393	354.58765	52.58446	0.02650	0.05039
E	100	131.95833	31.43714	372.21424	54.86994	0.04597	0.08378

- Example of a typical group:

Student ID	Group Work Attitude	Interest	Motivation	Self Confidence	Shyness	Level of Performance	Fluency in language	Score
1	2	1	1	1	2	1	1	9
2	2	3	3	2	1	2	2	15
3	2	2	2	2	1	1	2	12
4	3	1	1	2	2	2	1	12

ED = 14.93

GH = 6

Experiments

- Proof scalability
  Experiment with one data set with all 512 students' data
- Modifications
  - Applying 2-opt only for 20 % of the students/nodes (randomly selected)
  - Goal: Finding a good solution
  - Termination condition: stop after 200 iterations
- Result
  - CV values are higher than for the previous experiments with 100 students but still low (SD=0.37, CV=0.793)
  - → found stable, good solutions
- Comparison with an iterative algorithm
  - Average GH-Value: 4.2 (1.6)
  - Euclidean Distance: 2.49 (2.40)

Conclusion and Future Work

- Developed an approach to build heterogeneous groups
- Heterogeneity is based on
  - Different personality and performance attributes
  - A general measure of the goodness of heterogeneity
  - Coefficient of variation of goodness of heterogeneity values
- Implemented a tool that uses an ACO algorithm for optimization
- Experiments
  - Algorithm finds stable solutions close to the optimum with a data set of 100 students
  - Scalability was demonstrated with a data set of 512 students → algorithm found stable, good solutions
- Future Work
  - Combining the tool with an online learning system
  - Provide more options for user to adjust the algorithm