Seminar Paper: Optimized Grouping in Educational Activities

How to generate optimal groups of students algorithmically?

Lars Steffen März and Thomas Herde

February 26, 2023

TECHNISCHE UNIVERSITÄT DARMSTADT

Management Science/ Operations Research

Abstract

Grouping people to foster synergies has been an active field of research for decades, both in social sciences and management science. But also from the viewpoint of mathematical modelling and optimization, it is a challenging task due to its combinatorial and multi-objective nature. This paper focuses on the educational context by specifically considering group and topic assignments rooted in social preferences, topic interests, prior knowledge and other properties to an unprecedented extent. It provides highly customizable MIPs to model and `Java` code to solve such problems relying on the `Gurobi` library. Initial numerical tests, utilizing heuristically parameterized and randomly generated artificial user data, show computational tractability and provide approximately Pareto-efficient solutions to the user for moderate instance sizes of up to 30 students and 15 topics. Finally, the goal is to supply the chair for Management Science and Operations Research at TU Darmstadt with the developed software for practical application and evaluation.

Contents

1 Introduction and Problem Description

1.1 Task and Objectives

This paper is part of a Master's seminar of Prof. Dr. Weidinger and M.Sc. Constantin Wildt at the faculty for Operations Research and Management Science at the Technical University Darmstadt in the winter term 2022/23. The original task for the participants reads:

> *Dividing students into groups in educational activities subject to specific requirements is a challenging planning task. On the one hand, preferences regarding the topic to be dealt with must be taken into account. On the other hand, students are interested in forming teams with certain other people. On the part of the lecturer, it is important to ensure a balanced group size and a level of knowledge that is as evenly distributed as possible. Thus, it is a multi-criteria optimization problem in a pedagogical context. The goal of this seminar is to develop and formalize optimization approaches for the described problem and to implement a prototype making use of an optimization tool of the student's choice. The model, the concept behind it, and the prototype will be presented in a seminar presentation and a seminar paper.*

Hence, we identified three main groups of stakeholders with overlapping interests, for the sake of illustration assigned as follows:

- The participating students are generally interested in collaborating with certain other participants out of social motives. Simultaneously, they intend to deal with specific topics which might differ from the ones of their befriended peers.
- The professor and potentially her or his assistants aim at similarly and specifically sized groups as 'the size of a decision-making group influences what the group can achieve' [3] and member satisfaction [4], cf. Figure[1] 1. Besides this, a fair and balanced distribution of skills and knowledge among groups with a certain degree of diversity within groups might be aspirational since 'task-related diversity, was positively related to both quality and quantity of team performance' [6] and 'cognitive distance is a meaningful antecedent for group cognitive synergy' [7].
- The chair looks for an universal, versatile, cost-efficient, performant and easy to use software to compute groupings for future seminars. Therefore, only open-source or software with university license should be needed to run the code and the running-time should be well below of one day for moderate hardware. Operating the software shall not require prior experience in programming or mathematics and be intuitive regardless of seminar size or topic.

Ultimately, we aim to develop a practical software application, solving such problems by providing a sensible compromise between aforementioned goals. More precisely, we construct a mathematical optimization model of described problem and implement a solution algorithm to find optimal solutions to it, in order to support the grouping process.

User data input and solution output will work from `Excel`. All goals of the chair listed above will be fulfilled by all models below. Thus, they will not be mentioned explicitly going forward. Moreover, all of them will implement some form of group size limitation. Table 1 gives an overview over the models discussed in this paper and which of the remaining goals they address.

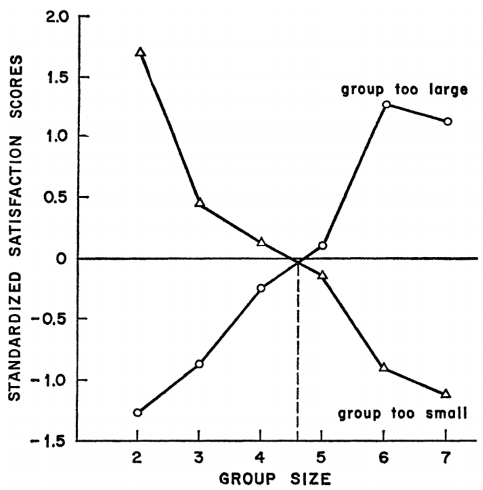


Figure 1: Reported satisfaction with group size (cf. [4], p. 48)

[1] The labels 'group too small' and 'group too large' tag the two graphs plotting how much (dis-)satisfaction resulted from the group being too small or large for the given group size, respectively. They do not refer to the abscissa, despite 4.6 members per group turning out to be empirically optimal.

Model Name	Problem	Social Sat.	Topic Sat.	Even Skill Distr.	Skill Divers.
Basic Naive	$(\mathcal{P}_1)$	✓			
Basic Improved	$(\mathcal{P}_2)$	✓			
Extend. f. Topic Assign.	$(\mathcal{P}_3)$	✓	✓		
Extend. f. Skill Distr.	$(\mathcal{P}_4)$	✓	✓	✓	✓
Efficient Frontier	$(\mathcal{P}_4)$	✓	✓	✓	✓

Table 1: Models and goals addressed

1.2 Notation

In the sequel, we will repeatedly make use of the following definitions.

$\mathbb{N} := \{1, 2, 3, ...\}$	$S :=$ number of students $\in \mathbb{N}$	$\mathcal{S} := [S] :=$ set of all students s
$\mathbb{N}_0 := \{0\} \cup \mathbb{N}$	$G :=$ number of groups $\in \mathbb{N}$	$\mathcal{G} := [G] :=$ set of all groups g
$[n] := \{1, ..., n\} \quad \forall n \in \mathbb{N}$	$T :=$ number of topics $\in \mathbb{N}$	$\mathcal{T} := [T] :=$ set of all topics t

Note that there is no loss of generality by identifying students, groups and topics with integers. It can be understood as an arbitrary enumeration of the real counterpart.

2 Basic Model

2.1 Naive Implementation

The following optimization problem was provided by Prof. Dr. Weidinger and M.Sc. Constantin Wildt as a starting point to our modelling journey.

$$\max_x \frac{1}{S} \sum_{s \in \mathcal{S}} \sum_{s' \neq s} q_{s,s'} x_{s,s'} \tag{$\mathcal{P}_1$}$$

$$\text{s.t. } \underline{m} - 1 \leq \sum_{s' \neq s} x_{s,s'} \leq \overline{m} - 1 \qquad \forall s \in \mathcal{S} \tag{1}$$

$$x_{s,s'} + x_{s',s''} \leq x_{s,s''} + 1 \qquad \forall (s, s', s'') \in \mathcal{S}^3 : \text{pairwise distinct} \tag{2}$$

$$x_{s,s'} = x_{s',s} \qquad \forall (s, s') \in \mathcal{S}^2 : s \neq s' \tag{3}$$

$$x_{s,s'} \in \{0, 1\} \qquad \forall (s, s') \in \mathcal{S}^2 : s \neq s'$$

It makes use of the following parameters that need to be provided by the user.

Param.	Range	Quantifier	Quantity	Description
$q_{s,s'}$	$\{0, 1\}$	$(s, s') \in \mathcal{S}^2 : s \neq s'$	$S(S-1)$	how much student s likes student s' ℓ_0-normalized
$\underline{m}$	$\mathbb{N}_0$		1	minimum group size
$\overline{m}$	$\mathbb{N}_0$		1	maximum group size

The below decision variables are to be optimized.

Variable	Range	Quantifier	Quantity	Description
$x_{s,s'}$	{0,1}	$(s, s') \in \mathcal{S}^2 : s \neq s'$	$S(S-1)$	one iff students s and s' share a group

The set of feasible solutions is restricted by these constraints.

Constr.	Quantifier	Quantity	Description
(1)	$s \in \mathcal{S}$	$2S$	bounds for group size
(2)	$(s, s', s'') \in \mathcal{S}^3$: pairw. dist.	$S(S-1)(S-2)$	transitivity of being in same group
(3)	$(s, s') \in \mathcal{S}^2 : s \neq s'$	$S(S-1)$	symmetry of being in same group

2.2 Improved Implementation

By removing the lower left half of the matrix $(x_{s,s'})_{s \neq s'}$ we save about half of the variables without loss of information due to its symmetry. The dominating number of transitivity constraints is also reduced by a factor of $1/2$ when remodelling as follows.

$$
\begin{aligned}
&\max_x \ \frac{1}{S} \sum_{s \in \mathcal{S}} \sum_{s' > s} (q_{s,s'} + q_{s',s}) x_{s,s'} && && (\mathcal{P}_2) \\
&\text{s.t. } \underline{m} - 1 \leq \sum_{s' > s} x_{s,s'} + \sum_{s' < s} x_{s',s} \leq \overline{m} - 1 && \forall s \in \mathcal{S} && (4) \\
&\quad x_{s,s'} + x_{s',s''} \leq x_{s,s''} + 1 && \forall (s,s',s'') \in \mathcal{S}^3 : s < s' < s'' && (5) \\
&\quad x_{s,s''} + x_{s',s''} \leq x_{s,s'} + 1 && \forall (s,s',s'') \in \mathcal{S}^3 : s < s' < s'' && (6) \\
&\quad x_{s,s'} + x_{s,s''} \leq x_{s',s''} + 1 && \forall (s,s',s'') \in \mathcal{S}^3 : s < s' < s'' && (7) \\
&\quad x_{s,s'} \in \{0,1\} && \forall (s,s') \in \mathcal{S}^2 : s < s'
\end{aligned}
$$

The parameters remain invariant.

Param.	Range	Quantifier	Quantity	Description
$q_{s,s'}$	$\{0,1\}$	$(s,s') \in \mathcal{S}^2 : s \neq s'$	$S(S-1)$	how much student s likes student s' ℓ_0-normalized
$\underline{m}$	$\mathbb{N}_0$		1	minimum group size
$\overline{m}$	$\mathbb{N}_0$		1	maximum group size

The amount of variables is halved as motivated above.

Variable	Range	Quantifier	Quantity	Description
$x_{s,s'}$	{0,1}	$(s,s') \in \mathcal{S}^2 : s < s'$	$S(S-1)/2$	one iff students s and s' share a group

Due to the reduction in decision variables we have to mind the ordering of indices in the transitivity constraints. Adding being a commutative operator, the ordering of summands does not need to be permuted. The symmetry constraint becomes redundant since the improved model is intrinsically symmetric by design.

Constr.	Quantifier	Quantity	Description
(4)	$s \in \mathcal{S}$	$2S$	bounds for group size
(5)	$(s,s',s'') \in \mathcal{S}^3 : s < s' < s''$	$S(S-1)(S-2)/6$	transitivity of being in same group
(6)	$(s,s',s'') \in \mathcal{S}^3 : s < s' < s''$	$S(S-1)(S-2)/6$	transitivity of being in same group
(7)	$(s,s',s'') \in \mathcal{S}^3 : s < s' < s''$	$S(S-1)(S-2)/6$	transitivity of being in same group

2.3 Performance Comparison

Hence the naive model has of order

$$2S + S(S-1)(S-2) + S(S-1) = S^3 + o(S^3)$$

many non-integrality constraints and about

$$S(S-1) = S^2 + o(S^2)$$

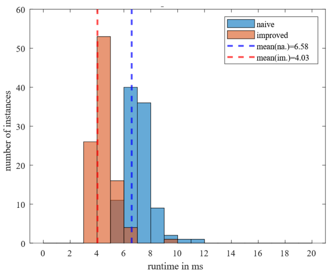
many binary variables. Whilst the improved model has of order

$$2S + \frac{3}{3!} S(S-1)(S-2) = \frac{1}{2} S^3 + o(S^3)$$

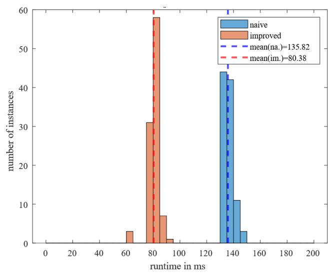
many non-integrality constraints and about

$$\frac{1}{2!} S(S-1) = \frac{1}{2} S^2 + o(S^2)$$

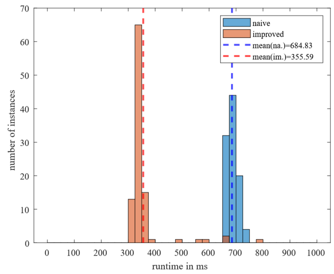
many binary variables. This represents an improvement of the leading Landau term by a factor of $1/2$ over the naive construction, respectively. As the numerical results in Figure 2 demonstrate, this improvement also translates into a running time speed up by almost the same factor across varying problem sizes and instances.



(a) $S = 10$



(b) $S = 30$



(c) $S = 50$

Figure 2: Runtime comparison of naive and improved implementation across $N = 100$ random instances

We not only notice an improvement regarding the average runtime but even an almost strict dominance of the improved version across the board. The few outliers might be attributable to a `Gurobi` heuristic or branch rule turning out to be unsuitable to those instances. The underlying distribution of the random instances and more information on the results are stated in Chapter 5.

3 Extended Model for Topic Assignment

3.1 Derivation

Building up upon the previous idea of social preferences to accommodate for peer groups and friends to work together, we now augment the functionality to also assign topics to groups rooted in topic interests of the participants as they are vital for providing a motivated work atmosphere. Concretely, besides the students $\mathcal{S}$, we explicitly model the set of groups $\mathcal{G}$ and additionally allow for an assignment of topics $\mathcal{T}$ to said groups.

In contrast to before, we now admit scalar social preferences $q_{s,s'} \in [-1, 1]$ for all $(s, s') \in \mathcal{S}^2 : s \neq s'$, as well as individual topic preferences $p_{s,t} \in [-1, 1]$ for all $(s, t) \in \mathcal{S} \times \mathcal{T}$. Beyond that, there might be students indifferent to either their group members or the topics offered, for instance if all participants are strangers or all topics seem equally satisfactory. In order to take account of this, we introduce parameters $\lambda_s \in [0, 1]$ to allow each student to split his or her ℓ_1-normalized vote (cf. Section 5.2 for explanation on vote normalization) between social and topic preferences as demonstrated in Figure 3.

To execute all assignments, we fundamentally change the model structure compared to $(\mathcal{P}_1)$ and $(\mathcal{P}_2)$. Namely, instead of an student-to-student approach, we now employ a three-layer student-to-group-to-topic approach. This means replacing the inter-social variables $x_{s,s'}$ by binary variables

$$x_{sg} := \begin{cases} 1 & \text{if person } s \text{ is in group } g, \\ 0 & \text{else} \end{cases} \qquad \forall (s, g) \in \mathcal{S} \times \mathcal{G}$$

to model the student-to-groups assignment and similarly

$$y_{gt} := \begin{cases} 1 & \text{if group } g \text{ has topic } t, \\ 0 & \text{else} \end{cases} \qquad \forall (g, t) \in \mathcal{G} \times \mathcal{T}.$$

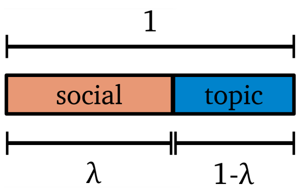


Figure 3: Vote split

to encode the group-to-topic assignment. Besides these explicit variables, we need two sets of implicit variables, derived from the ones above. Firstly, we need to decide whether two students s and s' are in the same group, i.e. if

$$\exists g \in \mathcal{G} : x_{sg} = x_{s'g} = 1. \tag{8}$$

For that, we use binary variables

$$a_{s,s',g} := \begin{cases} 1 & \text{if } s \text{ and } s' \text{ are both in group } g, \\ 0 & \text{else} \end{cases} \qquad \forall (s, s', g) \in \mathcal{S}^2 \times \mathcal{G} : s < s'$$

in constraint (14) to ensure

$$(8) \iff \sum_{g\in\mathcal{G}} a_{s,s',g} = 1$$

to hold. Secondly, for translating student-to-group and group-to-topic assignments into student-to-topic assignments, we again aim to model a condition of the form

$$\exists g \in \mathcal{G} : x_{sg} = y_{gt} = 1. \tag{9}$$

Demanding (15) gives

$$b_{sgt} := \begin{cases} 1 & \text{if } s \text{ has topic } t \text{ by being in group } g, \\ 0 & \text{else} \end{cases} \qquad \forall (s,g,t) \in \mathcal{S}\times\mathcal{G}\times\mathcal{T}.$$

and therefore by analogous arguments to before

$$(9) \iff \sum_{g\in\mathcal{G}} b_{sgt} = 1.$$

As an extension to the initial model, we want to reconsider the group size constraints. More specifically, some topics might be more labour-extensive than others. In recognition of this fact, the user may now enter topic-specific lower and upper bounds on the group size, i.e. integers $\underline{m}_t \leq \overline{m}_t$ for constraint (16).

Furthermore, in larger seminars, there might be the need to pass out the same topic to multiple groups. Or there could be some fundamental and therefore mandatory topics which are required as prior knowledge to understand the other ones. To implement such cases into our model, the user can bound the number of occurrences of topics via integral parameters $\underline{M}_t \leq \overline{M}_t$ in constraint (17).

3.2 Binary Program

The result of aforementioned modelling is the following linear binary program. It maximized the sum of the average social and topic satisfaction. As both summands are normalized on $[-1, 1]$, the objective value will be in $[-2, 2]$.

$$\max_{a,b,x,y} \frac{1}{S}\sum_{s\in\mathcal{S}}\sum_{s'>s}\left((\lambda_s q_{s,s'} + \lambda_{s'} q_{s',s})\sum_{g\in\mathcal{G}} a_{s,s',g}\right) + \frac{1}{S}\sum_{s\in\mathcal{S}}\sum_{t\in\mathcal{T}}\left((1-\lambda_s)p_{st}\sum_{g\in\mathcal{G}} b_{sgt}\right) \tag{$\mathcal{P}_3$}$$

$$\begin{aligned}
\text{s.t.}\quad & \sum_{g\in\mathcal{G}} x_{sg} = 1 && \forall s \in \mathcal{S} && (10)\\
& \sum_{t\in\mathcal{T}} y_{gt} \leq 1 && \forall g \in \mathcal{G} && (11)\\
& \sum_{t\in\mathcal{T}} y_{gt} \leq \sum_{s\in\mathcal{S}} x_{sg} && \forall g \in \mathcal{G} && (12)\\
& x_{sg} \leq \sum_{t\in\mathcal{T}} y_{gt} && \forall (s,g) \in \mathcal{S}\times\mathcal{G} && (13)\\
& x_{sg} + x_{s'g} - 1 \leq a_{s,s',g} \leq \frac{1}{2}x_{sg} + \frac{1}{2}x_{s'g} && \forall (s,s',g) \in \mathcal{S}^2\times\mathcal{G} : s < s' && (14)\\
& x_{sg} + y_{gt} - 1 \leq b_{sgt} \leq \frac{1}{2}x_{sg} + \frac{1}{2}y_{gt} && \forall (s,g,t) \in \mathcal{S}\times\mathcal{G}\times\mathcal{T} && (15)\\
& \underline{m}_t y_{gt} \leq \sum_{s\in\mathcal{S}} b_{sgt} \leq \overline{m}_t y_{gt} && \forall (g,t) \in \mathcal{G}\times\mathcal{T} && (16)\\
& \underline{M}_t \leq \sum_{g\in\mathcal{G}} y_{gt} \leq \overline{M}_t && \forall t \in \mathcal{T} && (17)\\
& a_{s,s',g} \in \{0,1\} && \forall (s,s',g) \in \mathcal{S}^2\times\mathcal{G} : s < s' &&\\
& b_{sgt} \in \{0,1\} && \forall (s,g,t) \in \mathcal{S}\times\mathcal{G}\times\mathcal{T} &&\\
& x_{sg} \in \{0,1\} && \forall (s,g) \in \mathcal{S}\times\mathcal{G} &&\\
& y_{gt} \in \{0,1\} && \forall (g,t) \in \mathcal{G}\times\mathcal{T} &&
\end{aligned}$$

This is a comprehensive list of all parameters used.

Param.	Range	Quantifier	Quantity	Description
$q_{s,s'}$	$[-1, 1]$	$(s, s') \in \mathcal{S}^2 : s \neq s'$	$S(S-1)$	how much student s likes student s' ℓ_1-normalized
p_{st}	$[-1, 1]$	$(s, t) \in \mathcal{S} \times \mathcal{T}$	ST	how much student s likes topic t ℓ_1-normalized
λ_s	$[0, 1]$	$s \in \mathcal{S}$	S	bias of s towards social over topic satisfaction
$\underline{m}_t$	$\mathbb{N}_0$	$t \in \mathcal{T}$	T	minimum group size for topic t
$\overline{m}_t$	$\mathbb{N}_0$	$t \in \mathcal{T}$	T	maximum group size for topic t
$\underline{M}_t$	$\mathbb{N}_0$	$t \in \mathcal{T}$	T	minimum number of groups with topic t
$\overline{M}_t$	$\mathbb{N}_0$	$t \in \mathcal{T}$	T	maximum number of groups with topic t

The decision variables x, y and their implied internal variables $a = a(x)$, $b = b(x, y)$ are presented here.

Variable	Range	Quantifier	Quantity	Description
$a_{s,s',g}$	{0,1}	$(s, s', g) \in \mathcal{S}^2 \times \mathcal{G} : s < s'$	$S(S-1)G/2$	one iff students s and s' share group g
b_{sgt}	{0,1}	$(s, g, t) \in \mathcal{S} \times \mathcal{G} \times \mathcal{T}$	SGT	one iff s is in group g having topic t
x_{sg}	{0,1}	$(s, g) \in \mathcal{S} \times \mathcal{G}$	SG	one iff s is in group g
y_{gt}	{0,1}	$(g, t) \in \mathcal{G} \times \mathcal{T}$	GT	one iff group g has topic t

Similarly, the below table gives an overview of all the employed constraints. Note that constraints (14) to (17) contain two inequalities each, resulting in double the quantity compared to the quantifiers.

Constr.	Quantifier	Quantity	Description
(10)	$s \in \mathcal{S}$	S	everyone is in exactly one group
(11)	$g \in \mathcal{G}$	G	every group has at most one topic
(12)	$g \in \mathcal{G}$	G	empty groups have no topic
(13)	$(s, g) \in \mathcal{S} \times \mathcal{G}$	SG	non-empty groups have a topic
(14)	$(s, s', g) \in \mathcal{S}^2 \times \mathcal{G} : s < s'$	$S(S-1)G$	$a_{s,s',g} = 1 \iff x_{sg} = x_{s'g} = 1$
(15)	$(s, g, t) \in \mathcal{S} \times \mathcal{G} \times \mathcal{T}$	$2SGT$	$b_{s,g,t} = 1 \iff x_{sg} = y_{gt} = 1$
(16)	$(g, t) \in \mathcal{G} \times \mathcal{T}$	$2GT$	bounds for group size
(17)	$t \in \mathcal{T}$	$2T$	bounds on topic occurrences

Supposing $1 \leq T \approx G \ll S < \infty$, problem $(\mathcal{P}_3)$ has of order

$$S + G + G + SG + \frac{2}{2!}S(S-1)G + 2SGT + 2GT + 2T = S^2G + o(S^2G)$$

many non-integrality constraints and about

$$\frac{1}{2!}S(S-1)G + SGT + SG + GT = \frac{1}{2}S^2G + o(S^2G)$$

many binary variables. For $G \leq S/2$, we asymptotically reduced the number of non-integrality constraints and only increased the asymptotic number of variables by a factor of G.

4 Extended Model for Skill Distribution

4.1 Derivation

As motivated in the introduction, we also want to promote skill diversity within groups and skill equality among groups. To this end, we distinguish between two types of personal characteristics: First, such skills $i \in \mathcal{I} := [I]$ that one either has or not and can therefore be represented by a binary variable

$$h_s^i \in \{0, 1\} \quad \forall (s, i) \in \mathcal{S} \times \mathcal{I}$$

with $h_s^i = 1$ if and only if student s possesses property i. For instance, one might encode whether or not the participants originate from a specific faculty, department or a set of them to ensure some degree of group interdisciplinarity. This might be useful since from experience, participants often look for fellow students they relate with from previous courses. However, this is often the case if they are enlisted in a similar major. In fact, this phenomenon could be observed at this seminar itself, with the participants splitting into a group of industrial engineering plus business informatics students and a group of mathematics majors. This led to vastly different focuses of the respective presentations, indicating that a more diverse group assignment might have yielded richer discussions during the group project stage. Therefore, the lower bound $\underline{h}_i$ or upper bound $\overline{h}_i$ for the number of students with property i per group could have been used to avoid such clustering. Secondly, there might also be properties $j \in \mathcal{J} := [J]$ that one can possess to some continuous extent, quantified by

$$e_s^j \in [0,1] \quad \forall (s,j) \in \mathcal{S} \times \mathcal{J}.$$

Such data might me useful for equal distribution of prior knowledge and relevant skills among different groups or intentionally uneven levels of leadership within the groups. For instance, at the chair for operations research and management science, it seems reasonable to pick a skill set similar to the choice on their thesis application homepage [8]. There, one is asked for a self-assessment regarding prior experience with optimization solvers, statistics software and general programming abilities. With these thoughts in mind, we augment Problem ($\mathcal{P}_3$) by technical hidden decision variables $\underline{c}_j$ and $\overline{c}_j$ to measure the lowest and highest cumulative group skill with respect to skill j.

4.2 Mixed-Integer Program

All of the aforementioned ideas are modelled in the subsequent inequalities.

$$\underline{h}_i \le \sum_{s\in\mathcal{S}} h_s^i x_{sg} + S(1-z_g) \qquad \forall (g,i) \in \mathcal{G} \times \mathcal{I} \tag{18}$$

$$\overline{h}_i \ge \sum_{s\in\mathcal{S}} h_s^i x_{sg} \qquad \forall (g,i) \in \mathcal{G} \times \mathcal{I} \tag{19}$$

$$z_g \ge x_{sg} \qquad \forall (s,g) \in \mathcal{S} \times \mathcal{G} \tag{20}$$

$$\underline{c}_j \le \sum_{s\in\mathcal{S}} e_s^j x_{sg} + S(1-z_g) \qquad \forall (g,j) \in \mathcal{G} \times \mathcal{J} \tag{21}$$

$$\overline{c}_j \ge \sum_{s\in\mathcal{S}} e_s^j x_{sg} \qquad \forall (g,j) \in \mathcal{G} \times \mathcal{J} \tag{22}$$

Note that adding S deactivates the constraints (18) and (21) for precisely the empty groups, i.e. where $z_g = 0$. The reasons behind empty groups are explained in answer $\mathbf{A_1}$ to question $\mathbf{Q_1}$ of Section 6.1. Summarizing, we added the following constraints.

Add. constr.	Quantifier	Quantity	Description
(18)	$(g,i) \in \mathcal{G} \times \mathcal{I}$	GI	lower hard skill bound for non-empty groups
(19)	$(g,i) \in \mathcal{G} \times \mathcal{I}$	GI	upper hard skill bound for groups
(20)	$(s,g) \in \mathcal{S} \times \mathcal{G}$	SG	$z_g = 1 \iff$ group g non-empty
(21)	$(g,j) \in \mathcal{G} \times \mathcal{J}$	GJ	lower exper. skill bound for non-empty groups
(22)	$(g,j) \in \mathcal{G} \times \mathcal{J}$	GJ	upper exper. skill bound for groups

Now, the high level idea to combine the objective of ($\mathcal{P}_3$) with the equality and diversity goals from this chapter is not to simply optimize a linear combination as all of the nuances, strength, weaknesses and trade-offs of different solutions are lost when condensing them into a single number. In fact, remembering the stakeholders and goals analysis from the introduction, we see an alignment of the topic and social objectives with the students' interests and the skill distribution objectives with the professor's interests. Hence, the one bailing out the other would not be a fair compromise. So, for the two parties concerned, we have two separate objectives.

$$\begin{aligned}
\max_{(x,y)} f_{\text{pref}}(x,y) &:= \frac{1}{S}\sum_{s\in\mathcal{S}}\sum_{s'>s}\left((\lambda_s q_{s,s'} + \lambda_{s'} q_{s',s})\sum_{g\in\mathcal{G}} a_{s,s',g}\right) + \frac{1}{S}\sum_{s\in\mathcal{S}}\sum_{t\in\mathcal{T}}\left((1-\lambda_s)p_{st}\sum_{g\in\mathcal{G}} b_{sgt}\right)\\
\max_{(x,y)} f_{\text{skill}}(x,y) &:= \sum_{j\in\mathcal{J}}\left(\sigma_j(\overline{c}_j - \underline{c}_j) + \mu_j \sum_{s\in\mathcal{S}}\sum_{s'>s}\left(\left|e_s^j - e_{s'}^j\right|\sum_{g\in\mathcal{G}} a_{s,s',g}\right)\right)
\end{aligned} \tag{23}$$

Note that for reasons of spacing, we omitted the variables $a = a(x)$, $b = b(x,y)$, $c = c(x,z)$ and $z = z(x)$ on the left-hand side of (23) since they are fully determined by x and y. Within f_{skill}, the equality and diversity interests are represented in a weighted manner using scalars σ_j and μ_j for all skills $j \in \mathcal{J}$. The former accomplishes this by punishing (or rewarding if one intends) large discrepancies between the lowest and highest cumulative group skill and the latter by rewarding (or punishing if one intends) individual skill differences between group members. In conclusion we added the following parameters

Add. param.	Range	Quantifier	Quantity	Description
h_s^i	$\{0,1\}$	$(s,i) \in \mathcal{S} \times \mathcal{I}$	SI	one iff student s has hard skill i
$\underline{h}_i$	$\mathbb{N}_0$	$i \in \mathcal{I}$	I	lower bound for hard skill i per group
$\overline{h}_i$	$\mathbb{N}_0$	$i \in \mathcal{I}$	I	upper bound for hard skill i per group
e_s^j	$[0,1]$	$(s,j) \in \mathcal{S} \times \mathcal{J}$	SJ	how much student s has exper. skill j
σ_j	$(-\infty,0]$	$j \in \mathcal{J}$	J	penalty for inequality w.r.t. group skill j
μ_j	$[0,\infty)$	$j \in \mathcal{J}$	J	reward for diversity within groups w.r.t. skill j

and decision variables.

Add. var.	Range	Quantifier	Quantity	Description
$\underline{c}_j$	$[0,\infty)$	$j \in \mathcal{J}$	J	lowest group competence w.r.t. exper. skill j
$\overline{c}_j$	$[0,\infty)$	$j \in \mathcal{J}$	J	highest group competence w.r.t. exper. skill j
z_g	{0,1}	$g \in \mathcal{G}$	G	one iff group g is non-empty

In total, we added $SG + 2GI + 2GJ = SG + 2G(I + J)$ many constraints and only G binary variables z, as well as $2J$ continuous variables c. Assuming $0 \leq I \approx J \ll T \approx G \ll S < \infty$, they add comparatively little complexity to the model despite transforming the IP ($\mathcal{P}_3$) to a MIP.
Now, the idea is to maximize both f_{pref} and f_{skill} simultaneously whilst still providing a couple of solution options - with key statistics attached to them - for the tutor to make an educated semi-automated decision. For the time being, we denote this idea via the non-rigorous notation

$$\max_{\substack{(a,b,c,x,y,z) \\ \text{s.t. (10),...,(22) hold}}} \Big(f_{\text{pref}}(a,b,x,y) \quad \text{and} \quad f_{\text{skill}}(a,b,c,x,y,z) \Big). \tag{$\mathcal{P}_4$}$$

To make this more rigorous, we first generalize the concept of optimization problems with multiple objectives by providing two essential definitions.

Definition. (Multi-Criteria Optimization Problem - cf. [2], p. 16)**.** *A multi-criteria optimization problem is a problem of optimizing multiple maps f_k from a set $\mathcal{K}$ on a single feasible set $\mathcal{X}$, i.e.*

$$\max_{x \in \mathcal{X}} \quad f_k(x) \quad \forall k \in \mathcal{K}. \tag{$\mathcal{MC}$}$$

Definition. (Pareto-dominated/efficient/optimal - cf. [5], p. 343)**.** *A feasible solution $x \in \mathcal{X}$ to a multi-criteria optimization problem ($\mathcal{MC}$) is called Pareto-dominated if there exists another feasible solution $x^* \in \mathcal{X}$ with*

$$f_k(x) \leq f_k(x^*) \quad \forall k \in \mathcal{K} \qquad \wedge \qquad \exists k' \in \mathcal{K} : f_{k'}(x) < f_{k'}(x^*).$$

Feasible solutions which are not Pareto-dominated are called Pareto-efficient or alternatively Pareto-optimal. The set of all Pareto-optimal solutions is called the efficient frontier.

Knowing what we are looking for, we will now derive two methods to find Pareto-optimal solutions.

4.3 Approximating the Efficient Frontier

Our goal is to locate Pareto-optimal points of problem ($\mathcal{P}_4$). However, finding all such points is computationally expensive, still matter of research [5] and does not yield much added value to the end user in comparison to a small collection of representative solutions. Hence, we try to design parameterized methods that yield such points and, if possible, distinct ones for varying parameters.

Blended Approach
The first approach optimizes a weighted sum of the objectives

$$\max_{x \in \mathcal{X}} \quad \sum_{k \in \mathcal{K}} w_k f_k(x) \tag{$\mathcal{MCB}[\mathcal{X}, w, \mathcal{K}]$}$$

depending on weights $\omega := (\omega_k) \geq 0$. Conveniently, `Gurobi` provides this option natively via the `ObjNWeight` attribute. By using different weights, we expect different solutions with diverse trade-offs.

Hierarchical Approach

For the second approach, we use `Gurobi`'s so-called `Multi-Criteria Degradation` environment to find said collection of solutions. For some general multi-criteria problem $(\mathcal{MC})$ it solves the sequence of problems

$$\begin{aligned} \max_{x \in \mathcal{X}} \quad & f_k(x) \\ \text{s.t.} \quad & f_{k'}(x) \geq f_{k'}(x^*_{k'}) - \tau_{k'} |f_{k'}(x^*_{k'})| \quad \forall k' \prec k \end{aligned} \qquad (\mathcal{MCH}[\mathcal{X}, \tau, \prec, k, \mathcal{K}])$$

for all $k \in \mathcal{K}$ in a given hierarchical order $\prec$ defined by `ObjNPriority`, e.g. $k' \in [k-1]$ for $\prec$ being $<$. Here, x^*_k denotes the optimal solution to $(\mathcal{MCH}[\mathcal{X}, \tau, \prec, k, \mathcal{K}])$ and $\tau := (\tau_k) \geq 0$ are tolerances to be specified by `ObjNRelTol`.

Concatenating both approaches and running them for different choices of weights ω, orderings $\prec$ and tolerances τ almost results in Algorithm 1. Namely, the next natural question is if all the solutions collected in $\mathcal{X}^*$ are actually Pareto-optimal. In fact, on their own, only the blended approach gives Pareto-optimal solutions whilst we need to put more thought into the hierarchical one: Since (w.r.t. $\prec$) later objectives only receive the optimal values of the previous maps as hard constraints, they are indifferent regarding the amount of reduction as long as it is within the tolerance. Consequently, free potential might be lost when multiple solutions provide the same $f_k(x^*_k)$ but different $f_{k'}(x^*_{k'})$ for $k' \prec k$, either by unfortunately choosing one with low $f_{k'}(x^*_{k'})$ values or by simply not optimizing otherwise untangled variables like $\underline{c}_j$ or $\overline{c}_j$ in the first place.

We solve this by fixing the optimal values after all objectives have been looped through in line 10 and asking the previous objective in reverse order if they wasted potential given the current $f_k(x^*_k)$ values. Having squeezed out all this otherwise lost head room, we finally add the solution to $\mathcal{X}^*$. That the above considerations are valid, is the statement of the following theorem.

Algorithm 1 –Efficient Frontier Approximation

1: //Initialization
2: $\mathcal{X} := \{(a, b, c, x, y, z)$ s.t. (10), ..., (22) hold$\}$
3: $\mathcal{K} := \{\text{pref}, \text{skill}\}$
4: $\mathcal{X}^* := \emptyset$
5:
6: //Hierarchical Approach
7: **for** different $\prec$ and $\tau \geq 0$ **do**
8:     **for** $k \in \mathcal{K}$ ascending in $\prec$ **do**
9:         solve $(\mathcal{MCH}[\mathcal{X}, \tau, \prec, k, \mathcal{K}])$ to get $(a, b, c, x, y, z)^*_k$
10:     **end for**
11:     **for** $k \in \mathcal{K}$ descending in $\prec$ **do**
12:         solve $(\mathcal{MCH}[\mathcal{X}, 0, \succ, k, \mathcal{K} \setminus \{\max_{\prec}(\mathcal{K})\}])$ to get $(a, b, c, x, y, z)^*_k$
13:     **end for**
14:     $\mathcal{X}^* := \mathcal{X}^* \cup (a, b, c, x, y, z)^*_{\min_{\prec}(\mathcal{K})}$
15: **end for**
16:
17: //Blended Approach
18: **for** different $w > 0$ **do**
19:     solve $(\mathcal{MCB}[\mathcal{X}, w, \mathcal{K}])$ to get $(a, b, c, x, y, z)^*$
20:     $\mathcal{X}^* := \mathcal{X}^* \cup (a, b, c, x, y, z)^*$
21: **end for**
22:
23: //Visualization
24: Plot $f_k(\mathcal{X}^*) \quad \forall k \in \mathcal{K}$

Theorem. *After running Algorithm 1, $\mathcal{X}^*$ only contains Pareto-optimal solutions.*

Proof. We have to distinguish between the hierarchical approach $(\mathcal{MCH}[\mathcal{X}, \tau, \prec, k, \mathcal{K}])$ in lines 7-15 and the blended approach $(\mathcal{MCB}[\mathcal{X}, w, \mathcal{K}])$ in lines 18-21.

- By contradiction, assume $(a,b,c,x,y,z)^* \in \mathcal{X}^*$ to be Pareto-dominated by some $(a,b,c,x,y,z) \in \mathcal{X}$. Then, due to $w_k > 0$ for all $k \in \mathcal{K}$, we see that $(a,b,c,x,y,z)^*$ would not be optimal for $(\mathcal{MCB}[\mathcal{X}, w, \mathcal{K}])$ in line 19 either. Hence $(a,b,c,x,y,z)^* \notin \mathcal{X}^*$ holds, yielding a contradiction.
- For contradiction, suppose $(a,b,c,x,y,z)^* \in \mathcal{X}^*$ to be Pareto-dominated by some $(a,b,c,x,y,z) \in \mathcal{X}$ and let $k \in \mathcal{K}$ be the largest index w.r.t. $\prec$ such that $f_k(a,b,c,x,y,z) > f_k((a,b,c,x,y,z)^*)$. Then, $(a,b,c,x,y,z)^*$ would not have been optimal for $(\mathcal{MCH}[\mathcal{X}, 0, \succ, k, \mathcal{K}]$ in line 12 and therefore been overwritten before reaching line 14. Consequently, we derive $(a,b,c,x,y,z)^* \notin \mathcal{X}^*$ again.

Since $\mathcal{X}^*$ is empty initially and only Pareto-efficient solutions get added, we have shown the claim. □

Our choice of weights ω, orderings $\prec$ and tolerances τ in lines 7 and 18 is shown in the next chapter. There, we will also take a look at the visual approximation of the efficient frontier created in line 24.

5 Numerical Tests

In this section, we take a look at some initial numerical tests. The underlying implementation can be accessed via `https://github.com/lsmaerz/optimized_grouping.git`. We start by explaining the distribution of the artificial user data before motivating a normalization to impede strategic voting. In the end, we state our success metrics utilized to assess the computed group and topic assignments.

5.1 Distribution Assumptions

For a precise description of our parameter distribution, we need some concepts from probability theory.

Definition. (Properties of Random Variables - c.f. [1]). (a) *A family $(\xi_i)_{i \in \mathcal{F}}$ of random variables (RVs) $\xi_i : (\Omega, \Sigma, \mathbb{P}) \to (\Xi, \Theta)$ is called identically distributed if it holds*

$$\mathbb{P}[\xi_i \in \theta] = \mathbb{P}[\xi_j \in \theta] \quad \forall \theta \in \Theta, \ i, j \in \mathcal{F}.$$

In this case, we write $\xi_i \sim \xi_j$ for all $i, j \in \mathcal{F}$.

(b) *A family $(\xi_i)_{i \in \mathcal{F}}$ of RVs $\xi_i : (\Omega, \Sigma, \mathbb{P}) \to (\Xi, \Theta)$ is said to be independent if*

$$\mathbb{P}[\xi_i \in \theta_i \quad \forall i \in \mathcal{S}] = \prod_{i \in \mathcal{S}} \mathbb{P}[\xi_i \in \theta_i] \quad \forall \mathcal{S} \subseteq \mathcal{F}, \ |\mathcal{S}| < \infty, \ \theta_i \in \Theta.$$

(c) *A family $(\xi_i)_{i \in \mathcal{F}}$ of RVs $\xi_i : (\Omega, \Sigma, \mathbb{P}) \to (\Xi, \Theta)$ is called independent, identically distributed (i.i.d.) if it satisfies (a) and (b).*

(d) *Let $\mathcal{U}[R]$ denote the uniform distribution for some measurable set R, especially $R \subseteq \mathbb{R}^n$ or $R \subseteq \mathbb{N}$.*

(e) *Define $\mathbb{1}_\Lambda : \Omega \to \{0,1\}$ with $\mathbb{1}_\Lambda(\omega) = 1$ if and only if $\omega \in \Lambda$ as the indicator function of a set $\Lambda \subseteq \Omega$.*

In the following, all RVs will be assumed to be independent and (pseudo) randomly generated with the generic java class `java.util.Random` in a seed-based approach. The seed being determined by the instance number, uniformity and comparability between the tested models is provided. The table below summarizes the main assumptions employed in our `NumericalTests` class. They have been validated to be realistic through Mr. Wildt's experience as scientific assistant from previous courses at TU Darmstadt.

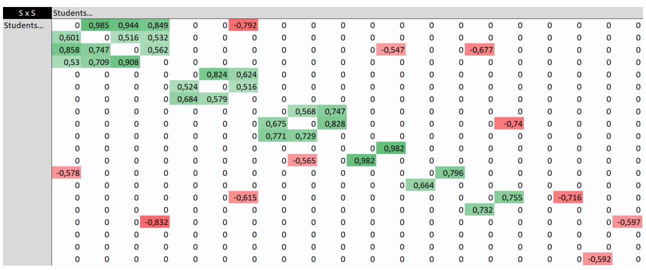
Param.	**Range**	**Quantifier**	**Distribution**
p_{st}	$[-1,1]$	$(s,t) \in \mathcal{S} \times \mathcal{T}$	$U_{st} \cdot \mathbb{1}_{[-1,-0.5) \cup [0.5,1)}(U_{st})$ for $U_{st} \sim \mathcal{U}[-1,1) \ \ \forall (s,t) \in \mathcal{S} \times \mathcal{T}$
λ_s	$[0,1]$	$s \in \mathcal{S}$	$\mathcal{U}[0,1)$
$\underline{m}_t$	$\mathbb{N}_0$	$t \in \mathcal{T}$	3
$\overline{m}_t$	$\mathbb{N}_0$	$t \in \mathcal{T}$	6
$\underline{M}_t$	$\mathbb{N}_0$	$t \in \mathcal{T}$	$\mathbb{1}_{[0,0.7/T)}(\mathcal{U}[0,1))$
$\overline{M}_t$	$\mathbb{N}_0$	$t \in \mathcal{T}$	$\underline{M}_t + 2$
h_s^i	$\{0,1\}$	$(s,i) \in \mathcal{S} \times \mathcal{I}$	$\mathbb{1}_{[0,p_i)}(\mathcal{U}[0,1])$ for $p_i \sim \mathcal{U}[0.5,0.6) \ \ \forall i \in \mathcal{I}$
$\underline{h}_i$	$\mathbb{N}_0$	$i \in \mathcal{I}$	$\mathcal{U}\{0,1\}$
$\overline{h}_i$	$\mathbb{N}_0$	$i \in \mathcal{I}$	$\mathcal{U}\{3,4,5\}$
e_s^j	$[0,1]$	$(s,j) \in \mathcal{S} \times \mathcal{J}$	$\mathcal{U}[0,1)$
σ_j	$(-\infty,0]$	$j \in \mathcal{J}$	$\mathcal{U}(-0.08,0]$
μ_j	$[0,\infty)$	$j \in \mathcal{J}$	$\mathcal{U}[0,0.05)$

The initialization of the matrix $(q_{s,s'})_{s \neq s'}$ is a process involving multiple steps that are stated in Algorithm 2. Intuitively speaking, it ensures that about 65% of participants already know somebody else prior to the course. Such students are clustered in peer groups of sizes two, three or four and like each other with $q_{s,s'} \sim \mathcal{U}[0.5, 1)$. Additionally, for every non-befriended ordered pair of students there is a 3% chance that the first feels averse towards the second. This is expressed in a negative social evaluation of $q_{s,s'} \sim \mathcal{U}(-1, 0.5]$.

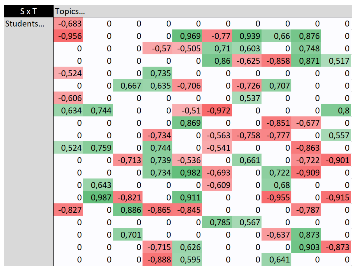
Algorithm 2 –Initialization of Social Preferences

```
//Parameters
hateProb := 0.03
socialShare := 0.65
maxPeerSize := 4
pos := 1

//Antipathy
for (s, s') ∈ S² : s ≠ s' do
    q_{s,s'} ~ U(−1, −0.5] · 𝟙_{[0,hateProb)}(U[0, 1))
end for

//Sympathy
while pos ≤ ⌊S · socialShare⌋ do
    groupSize ~ U{2, ..., maxPeerSize}
    for (s_1, s_2) ∈ {0, ..., groupSize − 1}² : s_1 ≠ s_2 do
        q_{pos+s_1, pos+s_2} ~ U[0.5, 1)
    end for
    pos := pos + groupSize
end while
```

Figure 4 illustrates a matrix instance constructed by Algorithm 2. It also depicts one exemplary initialization of the topic preferences.



(a) Matrix $(q_{s,s'})$



(b) Matrix (p_{st})

Figure 4: Instances of random matrices $(q_{s,s'})$ and (p_{st}) for $S = 20$ and $T = 10$

As for the orderings $\prec$, tolerances τ weights ω in Algorithm 1, we loop through both possible orderings, pick $\tau \in \{0.5, ..., 0.9\}$ and

$$\omega = (\omega_{\text{pref}}, \omega_{\text{skill}}) \in \{(1, 2^r) : r = 0, ..., 3\} \cup \{(2^r, 1) : r = 0, ..., 3\}.$$

In addition, we set $I := 2$ and $J := 3$ as motivated in Section 4.1. This sums up to 17 combinations for one frontier approximation, representing one 'Frontier' instance in Figure 5. Summarizing them across different problem instances instead, results in the numerical results for $(\mathcal{P}_4)$ called 'Multi' in Figure 5.

5.2 Normalization of User Input

Looking at the parameter choices above, we implicitly suppose all participants to provide honest topic and social preferences. However, this might not be case for individuals voting strategically. In that case, one might exaggerate preferences by setting the q-values to one for friends and minus one for everyone else.

To discourage such behaviour, we employ the following ℓ_1-normalizations.

$$\tilde{q}_{s,s'} \mapsto q_{s,s'} := \begin{cases} \frac{\tilde{q}_{s,s'}}{\sum_{s' \neq s} |\tilde{q}_{s,s'}|} & \text{if } \sum_{s' \neq s} |\tilde{q}_{s,s'}| > 0, \\ 0 & \text{else} \end{cases} \qquad \forall (s,s') \in \mathcal{S}^2 : s \neq s'$$

$$\tilde{p}_{st} \mapsto p_{st} := \begin{cases} \frac{\tilde{p}_{st}}{\sum_{t \in \mathcal{T}} |\tilde{p}_{st}|} & \text{if } \sum_{t \in \mathcal{T}} |\tilde{p}_{st}| > 0, \\ 0 & \text{else} \end{cases} \qquad \forall (s,t) \in \mathcal{S} \times \mathcal{T}.$$

So, more extreme ratings in absolute terms get encountered by antiproportional dilution effects. This results in everybody having one vote in absolute terms which is distributed according to λ_s between social and topic preferences. Also be aware that the basic models $(\mathcal{P}_1)$ and $(\mathcal{P}_2)$ were run with binary $q_{s,s'} \in \{0,1\}$ which were obtained by the translation

$$q_{s,s'} \mapsto \begin{cases} 1 & \text{if } (q_{s,s'})_{s' \neq s} \not\subseteq (-\infty, 0] \text{ and } s' = \max\left(\arg\max_i (q_{s,i})\right), \\ 0 & \text{else} \end{cases} \qquad \forall (s,s') \in \mathcal{S}^2 : s \neq s'.$$

However, for the numerical evaluation metrics, we assume the students to have continuous social preferences $q_{s,s'} \in [0,1]$ but being forced to make a binary decision by using the above method. Therefore, we will apply the success metric (24) to the assignment results computed by $(\mathcal{P}_1)$ and $(\mathcal{P}_2)$ on the basis of binary $q_{s,s'} \in \{0,1\}$. Doing so, we include possible information losses from the smaller user input range into our considerations.

5.3 Success Metrics

The four metrics considered aim to measure our success at reaching our introductory goals listed in Table 1. Note that for the two basic models $(\mathcal{P}_1)$ and $(\mathcal{P}_2)$, we equivalently replace $\sum_{g \in \mathcal{G}} a_{s,s',g}$ by $x_{s,s'}$ in the quantities below, respectively. We distinguish average social satisfaction

$$\frac{1}{S} \sum_{s \in \mathcal{S}} \sum_{s' > s} \left((q_{s,s'} + q_{s',s}) \sum_{g \in \mathcal{G}} a_{s,s',g} \right) \in [-1, 1], \tag{24}$$

average topic satisfaction

$$\frac{1}{S} \sum_{s \in \mathcal{S}} \sum_{t \in \mathcal{T}} \left(p_{st} \sum_{g \in \mathcal{G}} b_{sgt} \right) \in [-1, 1], \tag{25}$$

skill distribution gaps

$$\frac{\underline{c}_j}{\overline{c}_j} \in [0,1] \quad \forall j \in \mathcal{J}, \tag{26}$$

and skill diversity value

$$\frac{\sum_{s \in \mathcal{S}} \sum_{s' > s} \left(\left| e_s^j - e_{s'}^j \right| \sum_{g \in \mathcal{G}} a_{s,s',g} \right)}{\sum_{s \in \mathcal{S}} \sum_{s' > s} \sum_{g \in \mathcal{G}} a_{s,s',g}} \in [0,1] \quad \forall j \in \mathcal{J}. \tag{27}$$

Note that (24) and (25) measure the exact quantity maximized in $(\mathcal{P}_3)$ and $(\mathcal{P}_4)$ whilst (26) and (27) measure their respective goals differently to the objective in $(\mathcal{P}_4)$. There are multiple reasons for this:

- Both (26) and (27) are normalized to the unit interval, making comparisons between different solutions easier. Despite these advantages, we chose the linear representation in $(\mathcal{P}_4)$ to profit from the lower complexity of MILPs compared to MINLPs.
- In the case of the skill distribution gaps, the non-linear ratio in (26) again seems to be more telling than the difference $\overline{c}_j - \underline{c}_j$ since it disregards absolute levels of cumulative competence.

In contrast, the modifications done to obtain (27) solely serve normalization purposes as the quotient is still difficult to relate to a tangible real life quantity. At best, it can be interpreted as the achieved share of the maximum theoretic diversity. It reaches its theoretical maximum value of one if and only if $\left| e_s^j - e_{s'}^j \right| = 1$ for all group members of any group w.r.t. to skill j. Hence, as soon as a group contains more than two members, one cannot obtain a skill diversity value of one by the triangle inequality.

5.4 Numerical Results

We applied these metrics to randomly generated problem instances of $(\mathcal{P}_1)$ to $(\mathcal{P}_4)$ according to Sections 5.1 and 5.2. The tests ran on these hardware and software specifications:

- MIP-Solver: `Gurobi Optimizer version 10.0.0 build v10.0.0rc2 (win64)`
- CPU model: `AMD Ryzen 5 5600X 6-Core Processor, instr. set [SSE2|AVX|AVX2]`
- Thread count: `6 physical cores, 12 logical processors, using up to 12 threads`
- Coded in `Java (JDK 19)` on `Eclipse Editor 2022-12`

We made sure that `Gurobi` made use of all available threads. The following table gives an overview over the results.

Naive (N = 100)

	Runtime in ms	Social Satisfacion	Topic Satisfaction	Skill Gaps	Diversity Values
S = 10	7	0,74		0,62	0,34
S = 20	39	0,56		0,42	0,34
S = 30	136	0,51		0,36	0,33
S = 40	342	0,45		0,31	0,33
S = 50	685	0,42		0,27	0,33

Improved (N = 100)

	Runtime in ms	Social Satisfacion	Topic Satisfaction	Skill Gaps	Diversity Values
S = 10	4	0,74		0,62	0,34
S = 20	19	0,56		0,42	0,34
S = 30	80	0,51		0,35	0,33
S = 40	170	0,45		0,32	0,33
S = 50	352	0,42		0,27	0,33

Single (N = 100)

	Runtime in ms			Social Satisfacion			Topic Satisfaction			Skill Gaps			Diversity Values		
	T = 5	T = 10	T = 15	T = 5	T = 10	T = 15	T = 5	T = 10	T = 15	T = 5	T = 10	T = 15	T = 5	T = 10	T = 15
S = 10	96	190	271	0,74	0,75	0,76	0,18	0,10	0,08	0,59	0,59	0,58	0,33	0,33	0,33
S = 20	573	1.426	2.203	0,56	0,58	0,59	0,24	0,14	0,09	0,41	0,41	0,40	0,34	0,33	0,33
S = 30	2.060	5.474	8.805	0,51	0,53	0,54	0,25	0,15	0,10	0,34	0,32	0,32	0,34	0,34	0,33

Multi

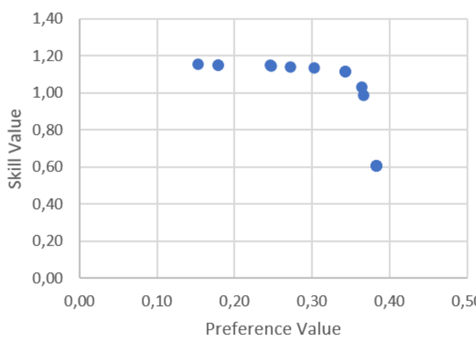
	Runtime in ms			Social Satisfacion			Topic Satisfaction			Skill Gaps			Diversity Values			
	T = 5	T = 10	T = 15	T = 5	T = 10	T = 15	T = 5	T = 10	T = 15	T = 5	T = 10	T = 15	T = 5	T = 10	T = 15	
S = 10	1.172	1.662	2.303	0,49	0,48	0,51	0,09	0,07	0,06	0,75	0,78	0,83	0,38	0,34	0,33	N = 85
S = 20	972.525	635.735	1.207.179	0,34	0,41	0,40	0,13	0,11	0,08	0,58	0,64	0,68	0,37	0,38	0,34	N = 34
S = 30	1.446.279	1.507.436	1.341.743	0,34	0,36	0,36	0,17	0,11	0,08	0,66	0,68	0,47	0,34	0,34	0,33	N = 17

Frontier

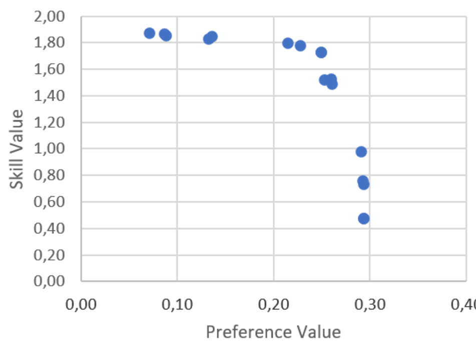
	Runtime in ms			
	T = 5	T = 10	T = 15	
S = 10	19.918	28.246	39.155	N = 5
S = 20	16.532.917	10.807.499	20.522.036	N = 2
S = 30	24.586.744	25.626.407	22.809.624	N = 1

Figure 5: Numerical results for Problems $(\mathcal{P}_1)$ to $(\mathcal{P}_4)$ and Algorithm 1 across N iterations

Moreover, the plots below showcase the efficient frontier approximations computed by Algorithm 1.



(a) $S = 20,\ \ T = 10,\ $ Seed $= 0$



(b) $S = 30,\ \ T = 10,\ $ Seed $= 1$

Figure 6: Efficient frontier approximations $(f_{\text{pref}}, f_{\text{skill}})(\mathcal{X}^*)$ from Algorithm 1

There are multiple interesting conclusions to be drawn from Figures 5 and 6:

- The runtime of all models grows in T and more dramatically in S which came to be expected when looking at the scaling of the number of variables and constraints w.r.t. to them. Additionally, each model takes much more time than its respective predecessor to an extent where it throws off the conditional coloring which is why we omitted it for $(\mathcal{P}_4)$ and Algorithm 1. However, all instances of all models terminated after at most 7.12 hours, as demanded in the introduction.

- The social satisfaction consistently falls as S increases. This arguably comes from our distribution assumptions on $(q_{s,s'})$ in Algorithm 2. There, we established everybody to have between zero and three friends and a probability of 3% to dislike each of the others. Consequently, the number of people a given student has no preference to or even hates, grows proportionally in S whilst the number of peers remains bounded above in a constant manner.
- The topic satisfaction appears to correlate positively with the ratio S/T. This is possibly rooted in our assumption that social preferences are clustered in peer groups but topic preferences do not correlate with them. Hence, simply having less topics to choose from causes a larger topic preference overlap within friend groups.
- The skill gap values, quantifying distributional fairness among groups, shrink even faster with S than the social satisfaction. This can be attributed to the increasing number of groups, seemingly making it harder to keep all of them around an average cumulative skill level. Penalizing unfair skill gaps dampens this reduction when comparing $(\mathcal{P}_1)$ to $(\mathcal{P}_3)$ with $(\mathcal{P}_4)$.
- Conversely, diversity must be rewarded to be increased as the same comparison as before shows.
- These equality and diversity benefits come at the cost of reduced social and topic happiness, as comparing $(\mathcal{P}_1)$ to $(\mathcal{P}_3)$ with $(\mathcal{P}_4)$ again, indicates.

The last three bullet points insinuate that our extensions to obtain $(\mathcal{P}_4)$ are to some degree effective in terms of finding a compromise between improving equality and diversity while keeping social and topic preferences. The precise trade-off is depended on the parameter choices made.

5.5 Problems During Testing

To obtain the previously discussed results, we had to overcome numerous problems over the course of our numerical tests. The three most significant ones and our counter measures are explained here.

Infeasible Random Instances
The distribution assumptions from Section 5.1 sometimes yielded infeasible instances since there was no assignment obeying the given bounds on the size (16) and number of occurrences (17) whilst simultaneously fulfilling the hard skill constraints (18) and (19). We fixed this by slightly loosening the parameters in the aforementioned constraints and slightly modifying instances with extreme distributions in h_s^i by increasing or reducing the number of students with $h_s^i = 1$.

Tuning of Parameters
During testing, we faced to problems described in $\mathbf{Q_4}$ of Section 6.1 regarding σ_j and μ_j. We coped with them by a simple trial-and-error procedure, i.e. iteratively solving instances, analyzing success metrics on the efficient frontier and adapting the parameters until reaching satisfactory trade-offs in the results as described in $\mathbf{A_4}$.

Long Runtime for Frontier Approximation
As we recognized in Section 5.4, the runtimes for $(\mathcal{P}_4)$ as stand-alone and as part of Algorithm 1 explode even for moderate S and T. Tracking the computation process revealed that the majority of that time gets spend on the third and therefore last step of the pyramid approach of Algorithm 1. The console output showed that this problem has very few feasible solutions as the non-degradation constraint (due to $\tau_k = 0$) from the previous step is very restricting. As a consequence, `Gurobi` heuristically finds good solutions quickly and fails to improve them significantly later. Instead, mostly the dual bound is reduced while the primal one stays still.

As counter measures, we set time limits proportional to the maximum running time provided by the user for each of the three pyramid steps and forced `Gubrobi` to solely focus on finding better feasible solutions after 65% of the total time has elapsed by setting the `ImproveStartTime` parameter accordingly. Additionally, we vastly increased the duality gap to 10% as the dual bound remained high due to the strategy change. Despite these radical measures, the efficient frontier approximation remained in large parts convex and smooth as can be seen in Figure 6. Asking `Gurobi` for other tuning suggestions via the `tune` method, resulted in different recommendations for different instances. Since none of them provided significant or universal improvements, we did not incorporate them permanently into the final code.

6 Discussion and Outlook

In this chapter, we want to answer possible questions that might occur and take a step back to consider which modelling choices we made that deserve reconsideration.

6.1 Discussion

$\mathbf{Q_1}$ Why not merge topics and groups to reduce models ($\mathcal{P}_3$) and ($\mathcal{P}_4$), i.e. why does the intermediate group layer exists and how is G chosen?

$\mathbf{A_1}$ As motivated in Section 3.1, certain topics might require a larger head count than others. This was the reason we introduced topic-specific bounds on the group size via integers $\underline{m}_t$ and $\overline{m}_t$. Consolidating groups and topics would not allow for that. In analogy, we gave reasons why setting a minimum or maximum number of topic occurrences could be useful. A simplified two-layer model would actually allow for replicating $\underline{M}_t$, by creating as many groups with that topic and making them mandatory via $z_g = 1$, and also $\overline{M}_t$ by the same idea but not externally setting z_g to some value. Omitting the middle layer would also spare the need for calculating G which is done by $G := S/\min_{t\in\mathcal{T}} \underline{m}_t$ and generally causes excessive groups to exists. So, simplifying the model is a reasonable idea yielding a smaller model but with less customization options.

$\mathbf{Q_2}$ How to gather student data without subconsciously influencing the professor's evaluation of their assignments?

$\mathbf{A_2}$ One way to enable anonymous voting is via pseudonymized ballots. More concrete, the professor hands out ballots with unique codes on them. After voting and calculating, the professor publishes the results using the pseudonyms. Hence, all students know their respective group and topic via their code they remembered. A more frictionless approach without anonymity would be to organize voting via a software like `Moodle`, allowing a direct export of the results into `Excel` where the models read in the parameters. Assuming the professor to be neutral, we advise using the latter for simplicity reasons.

$\mathbf{Q_3}$ Does showing the optimization problem to the students enable strategic voting?

$\mathbf{A_3}$ Via the ℓ_1-normalization technique of Section 5.2, we prevent strategic voting by the means of exaggeration. However, showing Problem ($\mathcal{P}_4$) to the participants might allow them to deduce a different manipulation idea: Peer group collusion, i.e. aligning social and topic preferences within cliques, inclines the model to cluster them into the same group. Additionally and even more undesirable, individuals could intentionally underestimate their skills e_s^j to promote solutions where they have more experienced group members since the model is rewarded for that twofold: On the one hand for cumulative skill equality and on the other hand via relative skill diversity between the alleged unskilled and skilled. Consequently, we recommend keeping the optimization problem largely private.

$\mathbf{Q_4}$ What are appropriate choices for the parameters σ_j and μ_j in f_{skill}?

$\mathbf{A_4}$ To motivate this question, it is helpful to revise the definition f_{skill} in (23) in contrast to f_{pref}. We see that in f_{pref} one has a natural normalization $1/S$ of both summands and an individual weighting λ_s between them. Neither of these two properties hold for f_{skill} since - as discussed in Section 5.3 - there is no obvious linear normalization in the objective and therefore the professors choice of σ_j and μ_j is not trivial. Beyond that, the term punishing inequality scales less dramatic in the average group sizes than the diversity term where each added group member provides $\left|e_s^j - e_{s'}^j\right|$ values relative to all the existing members. So one cannot simply copy parameter choices from one seminar to another. All in all, problem instances are often alike w.r.t. desired group sizes and relevance of equality versus diversity. Therefore, our tuned choices in Section 5.1 provide a rough magnitude and good starting point for own tuning dedicated to a particular instance.

$\mathbf{Q_5}$ Which model to use in which use case?

$\mathbf{A_5}$ Building upon the previous answer, we clearly recommend Problem ($\mathcal{P}_4$) to users with some amount of experience and patience for tuning σ_j and μ_j, for picking a solution from the efficient frontier approximation and also for the runtime itself. If diversity and equality are particularly relevant, this method is without alternatives as our numerical results demonstrated. On the flip side, models ($\mathcal{P}_2$)

and ($\mathcal{P}_3$) are more straightforward. As runtimes remain below ten seconds either way, we universally recommend using ($\mathcal{P}_3$) even if there are no topics to be assigned. In that case, providing one dummy topic and $\lambda_s = 1$ for all $s \in \mathcal{S}$ boils the model down to the functionality of the basic one whilst still enabling granular social preferences $q_{s,s'} \in [0, 1]$.

$\mathbf{Q_6}$ How does the concrete workflow look like when applying the algorithms?

$\mathbf{A_6}$ The user enters all parameter choices into a preformatted `Excel` file. In the `Java` class `StartHere`, one inserts the corresponding source path, destination path and chooses the model to be solved. Then, a new `Excel` file with the results will be generated at the specified location. For running your own numerical tests or checking the existing ones, we refer to the `Java` class `NumericalTests`. Figure 7 visualizes this process and the involved software.

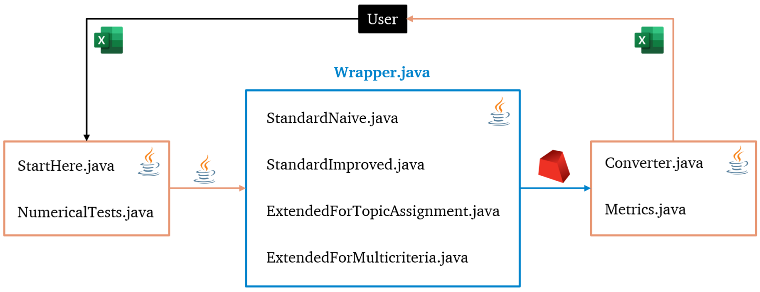


Figure 7: Schematized workflow for application and testing

6.2 Outlook

Beyond these questions, there are many other worth investigating in the future, among them:

$\mathbf{Q_7}$ How accurate are students' self-assessments regarding prior experience e_j in ($\mathcal{P}_4$)? Being a deeply psychological question, further examining of pertinent literature is necessary.

$\mathbf{Q_8}$ How sensitive is the running time w.r.t. I and J? Preliminary results show a low dependence but more thorough testing is necessary for validation.

$\mathbf{Q_9}$ Are there other important proxies of group performance that cannot be expressed appropriately in the model frameworks we developed?

$\mathbf{Q_{10}}$ Does modifying f_{skill} in (23) to

$$f_{\text{skill}}(x, y) := \sum_{j \in \mathcal{J}} \left(\sigma_j (\overline{c}_j - \underline{c}_j) + \mu_j \sum_{s \in \mathcal{S}} \sum_{s' > s} \left(\left| e_s^j - e_{s'}^j \right|^r \sum_{g \in \mathcal{G}} a_{s,s',g} \right) \right)$$

for some $r > 1$ yield different success metric values, especially diversity values?

$\mathbf{Q_{11}}$ Are there more telling or appropriate ways to measure skill equality and diversity, ideally in a linear and possibly normalized manner?

$\mathbf{Q_{12}}$ Does applying our software in a real case yield satisfactory results for all stakeholders mentioned in the introduction?

Despite surpassing the scope of this paper, all the above open questions deserve additional attention as their answers might give additional insights and allow for prospective model improvements.

7 Conclusion

In this paper, we considered the problem of assigning participants into groups and topics to said groups in an educational context. Doing so, we took social preferences, topic interests, prior experience, skills and other properties into account. We strived for a maximization of social and topic satisfaction whilst ensuring that the skills and experiences are distributed equally among and diverse within groups. To this end, we constructed a sequence of MIPs ($\mathcal{P}_1$) to ($\mathcal{P}_4$) of increasing functionality and customizability leading to Algorithm 1 which approximates the efficient frontier of described multi-criteria problem. This gives the user a concise collection of efficient trade-offs between the objectives entered.

We kept user friendliness in mind by keeping the majority of the user data input work within instructive `Excel` sheets whilst all computations are managed via `Java` communicating with the `Gurobi` solver. This reduces the user interaction with the programming code to a minimum. Initial numerical tests, utilizing heuristically parameterized and randomly generated artificial user data, suggest computational tractability and provide approximately Pareto-efficient solutions for moderate instance sizes. Practical testing and evaluation of stakeholder feedback remains subject to the implementation of our software in a real world educational environment.

A pre-recording of the accompanying seminar presentation to this paper can be accessed via `https://youtu.be/xTif4Ucr4Cg` with additional details on Chapter 6 under `https://youtu.be/II4A_9FQLZY`.

References

[1] Volker Betz. *Probability Theory (Lecture Notes)*. TU Darmstadt. Winter term 2020/21.

[2] Matthias Ehrgott. *Multicriteria Optimization. Second Edition*. 2005. ISBN: 3-540-21398-8.

[3] Nicolas Fay, Simon Garrod, and Jean Carletta. *Group Discussion as Interactive Dialogue or as Serial Monologue: The Influence of Group Size*. American Psychological Society. 2000. DOI: `10.1111/1467-9280.00292`.

[4] Richard J. Hackman and Neil Vidmar. *Effects of Size and Task Type on Group Performance and Member Reactions*. American Sociological Association. 1970. DOI: `10.2307/2786271`.

[5] Pascal Halffmann et al. *Exact Algorithms for Multiobjective Linear Optimization Problems with Integer Variables: A State of the Art Survey*. 2022. DOI: `10.1002/mcda.1780`.

[6] Sujin K. Horwitz and Irwin B. Horwitz. *The Effects of Team Diversity on Team Outcomes: A Meta-Analytic Review of Team Demography*. Southern Management Association. 2007. DOI: `10.1177/0149206307308587`.

[7] Nicoleta Meslec and Petru L. Curşeu. *Too Close or Too Far Hurts: Cognitive Distance and Group Cognitive Synergy*. 2013. DOI: `10.1177/1046496413491988`.

[8] Chair for Operations Research and Management Science at Department 1. *Bewerbungsformular für Abschlussarbeiten am Fachgebiet Management Science / Operations Research*. TU Darmstadt. 2023. URL: `https://www.or.wi.tu-darmstadt.de/bwl3_lehre/bwl3_abschlussarbeiten/bwl3_bewerbungsformular/bwl3_abschlussarbeitsformular.de.jsp`.