Tractable Model for the Optimization of Group Assignments Based on Preferences Over Positions and Other Individuals

Roope Pantsar Hong Ming Tan Jussi Keppo

June 2023

Abstract

This paper proposes an optimization model for personalized role and group assignment, tailored to individuals' preferences for colleagues and roles. Our model gives a solution that maximizes the overall satisfaction of the individuals and overcomes the limitations of traditional matching by allowing for a broader range of individual preferences and role requirements. The model has broad applications, such as assigning pilots to flights in aviation while considering both flight preferences and crew compatibility, a use case for which we created and tested a mobile application.

1 Introduction

The modern era demands innovative solutions for optimization problems in various domains. With the increasing complexity of tasks and diversification of individual skill sets, the need to optimally match people to roles within groups becomes a crucial factor in many environments. This task, however, is far from trivial; it demands sophisticated computational models and algorithms that take into account multiple dimensions and preferences to arrive at the most beneficial assignments. In this context, we propose an advanced optimization model and have developed a companion mobile application, specifically targeted for pilot scheduling. This mobile application gathers preferences from pilots and uses our model to optimally schedule them for flights.

Traditional optimization algorithms largely focus on aligning individuals with roles based on their qualifications, skills, or compatibility with the role requirements. While this is undoubtedly important, it overlooks a significant dimension of team dynamics and performance – the interpersonal relationships and preferences of team members. Individuals tend to perform better when they work with colleagues whom they prefer or feel comfortable with. Therefore, a comprehensive optimization solution should take into account not just role suitability, but also the personal preferences of individuals regarding their colleagues.

The algorithm we propose in this paper, and which is utilized in our mobile application, is designed to address this gap. By considering both role suitability and personal preferences, our model enables a more holistic optimization process. It involves a comprehensive set of variables, constraints, and parameters, handling different dimensions of the problem with a rigorous mathematical framework. It assigns individuals to roles and groups in such a way that optimally aligns with their skillsets and role preferences while also ensuring they are grouped with preferred colleagues to the best extent possible.

Our algorithm is grounded in the principles of mathematical optimization, featuring variables representing individual assignments to groups and roles, as well as the interrelation of these assignments with their preferred colleagues. The constraints of the model ensure realistic and feasible assignments that adhere to the capacities and limitations of roles and groups, and individuals' abilities to take on multiple roles or join multiple groups. Furthermore, the model is parameterized to capture personal preferences in terms of desired colleagues and preferred roles. A unique feature of our model is the incorporation of a parameter, δ_i, that balances an individual's relative care about their role assignments and the colleagues they are assigned with.

The problem that our algorithm addresses is pervasive and applicable across various domains. Consider a corporate setting where project teams need to be assembled, taking into account both the employees' competencies and their interpersonal dynamics. Similarly, in the context of education, forming study groups that consider both students' academic interests and their preferred partners can lead to more effective learning outcomes. In aviation, the mobile application we developed uses our model to optimally schedule pilots for flights by considering both their role preferences and preferred colleagues, which can enhance operational efficiency, job satisfaction, and even flight safety.

We anticipate that this innovative approach will be instrumental in yielding better-performing groups across a myriad of domains, fostering a more harmonious and productive environment. As the world becomes more interconnected, the importance of considering personal preferences alongside role suitability in group and role assignments can only increase. Thus, we believe that our proposed optimization model, in conjunction with the mobile application for pilot scheduling, will contribute significantly to this growing field and provide a powerful tool for improving group performance and satisfaction.

2 Literature Review

The foundation for the model proposed in this paper can be traced back to seminal works in the field of combinatorial and mathematical optimization. Pioneering studies that motivated this research are the stable marriage problem proposed by Gale and Shapley in 1962 and the stable roommate problem advanced by Irving in 1985.

The Stable Marriage Problem (SMP) introduced by Gale and Shapley (1962) presents a solution for a two-sided matching market, where each participant on one side of the market has a preference ranking for the participants on the other side. The solution seeks a stable matching, i.e., a matching where there is no pair of participants who would prefer each other over their current partners.

Following the SMP, Irving's Stable Roommate Problem (SRP) Irving (1985) took a similar preference-based approach but in a single-sided market, where participants are matched in pairs, with each having a preference ranking over all others. The SRP further emphasized the importance of individual preferences, mirroring the concept of personal preferences for group compositions in our proposed model.

While these classical problems provided a robust starting point for our investigation, the focus on stability in these models limited their applicability to more complex real-world scenarios where multiple conflicting objectives and various constraints are present. The transition from stability to optimization was a necessary step towards addressing these complexities. Unlike stable matching, optimization allows for more flexible and comprehensive solutions by enabling the consideration of multiple criteria and the balancing of various trade-offs.

In addition, many similar matching theory algorithms, while working in polynomial time, require each individual to have a list of strict preferences which for the purposes of the application of this specific problem, are difficult to obtain accurately. Other papers like Dutta and Massó (1997) explain a modified version of the deferred acceptance algorithm, first proposed by Gale and

Shapley (1962) to find stable assignments of individuals in firms while incorporating an individual's preferences for others. Other works include different parts of workings of the model in matching terms such as Cechlárová and Ferková (2004) which includes the addition of roles to the algorithm, allowing an individual to choose which role they would prefer another individual to take in the matching assignment. While this choice is not something that our optimization provides, choosing someone else to take one out of 2 roles, such as in the example of pilot scheduling, implies that one prefers themselves to take the other role and to be matched with that specific individual, which is a preference that is accounted for in our optimization.

The other type of algorithms that are similar to our model are the ones that use optimization to create groups. However, the key differences between these methodologies is that they usually prefer to use more objective and observable data in their formulations for the objective functions such as Rabiei et al. (2023) which uses the similar types of binary variables to assign volunteers in emergencies, and uses a mathematical formula considering distance to the Emergency Department(ED), the workload required in the assigned ED, and the importance of difficult situations encountered by the volunteer to calculate and minimize the degree of unsatisfied preferences of volunteers. Another is Yekta et al. (2023) which uses optimization to maximize the objective value of individuals in a matching, also using a multi-objective approach, however the other objective relies of minimizing the matching theory concept of coalitions of individuals who would change their assignments to become better off than the matching that they are provided. Another interesting similarity is between our model and Lin et al. (2010) in which students are matched to groups by their capability in certain disciplines. The methodology for this optimization is vastly different, but a similarity is that it includes the necessity that at least one individual in the group is interested in the discipline that is chosen for that group. This is implemented differently as the rankings and the interests are based on the results of test scores that are given to the algorithm before anything else, but the way the groups are matched with binary variables and the fact that what are perceived to be preferences are involved link this to our algorithm.

The academic landscape of assignment matching has undergone significant evolution, starting from strict preference-based stable matching models to contemporary optimization algorithms. The former have contributed substantially to our understanding of preference-based assignments, as exemplified by the seminal works of Gale-Shapley and Irving. They provided insightful solutions with a focus on stability, ensuring no pair of individuals would prefer to swap their current assignments. However, these models tend to operate on rigid preference lists and, while offering stable solutions, may not always lead to overall satisfaction. This caveat becomes more apparent in real-world scenarios where preferences can be nuanced and multifaceted, extending beyond binary choices.

On the other hand, optimization algorithms offer a different perspective. They excel at handling complex, instance-specific scenarios and are highly efficient at deriving optimal solutions based on the data provided. However, a significant shortfall of many existing optimization models is their lack of direct consideration for individual preferences. Rather, these models often infer preferences indirectly from other data, which may not reflect the true preferences of the individuals involved.

From this standpoint, our proposed model seeks to bridge this gap by blending the merits of both stability and optimization. It leverages the power of optimization algorithms for handling complex, multifaceted problems while directly incorporating individual preferences, akin to the early models of stable matching with the additional benefit of considering an individual's preferences for roles and groups as well as the individuals that they will be in those groups with. Our model embodies the evolution of assignment matching from simplistic, rigid models to one that is flexible, personalized, and capable of managing complex real-world scenarios.

3 Simple Model

Consider an instance of this model which includes six people, 1-6 who are being assigned to groups 1 and 2, where group 1 contains roles 1 and 2, group 2 contains role 3. The roles 1-3 have minimum sizes 1, 2, 2 and maximum sizes 2, 3, 3, respectively. The groups' maximum and minimum sizes are found by adding the minimum and maximum sizes of the roles in each, so the minimum and maximum sizes of group 1 (containing roles 1 and 2) would be 3 and 5, respectively, and the minimum and maximum sizes of group 2 (containing role 3) would be 2 and 3, respectively. In addition, in this example we only allow each person to be a part of one role, though this is not a necessary constraint. Each individual i's preferences for group members are shown below, labelled $\mathcal{P}_i$ and containing the indexes of said individual's preferred group members.

$$\mathcal{P}_1 = \{5,6\}, \quad \mathcal{P}_2 = \{5\}, \quad \mathcal{P}_3 = \{1\}, \quad \mathcal{P}_4 = \{5\}, \quad \mathcal{P}_5 = \{2,4\}, \quad \mathcal{P}_6 = \{1\}$$

The individuals also have preferences for the roles that they would like to be in, labelled $\mathcal{R}_i$ for each individual i which contain the indexes of the roles that they would like to be assigned to. The preferences for this example are shown below.

$$\mathcal{R}_1 = \{1\}, \quad \mathcal{R}_2 = \{1\}, \quad \mathcal{R}_3 = \{2\}, \quad \mathcal{R}_4 = \{2\}, \quad \mathcal{R}_5 = \{3\}, \quad \mathcal{R}_6 = \{3\}$$

The results of the optimization while taking only into account the interpersonal preferences, role preferences, and both together are shown below. As can be seen when an algorithm focuses on only one of these facets when creating teams, they are often able to increase that objective to a high value, but at the expense some other desirable qualities. In the case of this model, when either the preferred members or the preferred roles are maximized by themselves, the other sees a large loss.

However, this model can also optimize the two together, and as seen in this small example, both types of preference can be increased without incurring a large loss to the other.

	Total Preferred Members Assigned	**Total Preferred Roles Assigned**	***i* in Role 1**	***i* in Role 2**	***i* in Role 3**
Only $\mathcal{P}_i$	7	1	2	4, 5	1, 3, 6
Only $\mathcal{R}_i$	0	6	1, 2	3, 4	5, 6
Both $\mathcal{R}_i$ and $\mathcal{P}_i$	7	3	1	3, 6	2, 4, 5

4 Explanation of Model Construction

4.1 Sets

In the context of our optimization model, we navigate three main sets: People ($\mathcal{P}$), Roles ($\mathcal{R}$), and Groups ($\mathcal{G}$), where their respective cardinalities are denoted as N, T, and M. The primary objective of the model revolves around formulating an efficient mechanism for assigning people to specific roles, where each role is uniquely linked to a group.

The first set, $\mathcal{P}$, encompasses N unique individuals who are candidates for assignment to different roles within specific groups. Each person in $\mathcal{P}$ is a distinct entity with varying characteristics such as skills, interests, and availability. However, our model presumes a generalizable approach, where no such attributes constrain the assignment of an individual to a certain role.

The second set, $\mathcal{R}$, constitutes T distinct roles. These roles delineate specific tasks or functions that are essential within each group. These roles can range from specific titles like ‘Product Manager’, ‘Software Architect’, ‘Sales Consultant’, to more general ones such as ‘Team Member’. Each role in $\mathcal{R}$ is distinct and is uniquely associated with one group, ensuring a partition of the total set of roles across the different groups. However, a single role can be filled by multiple individuals from the $\mathcal{P}$ set if the dynamics of the group allow for it.

The third set, $\mathcal{G}$, comprises M unique groups. Each group represents an organizational entity like a team or a division, requiring the assignment of individuals to specific roles to ensure its functionality. Each group has a unique set of roles derived from the partitioned set $\mathcal{R}$. While a group is linked to unique roles, it does allow for multiple instances of individuals filling the same role, catering to the flexibility of the model.

This unique partitioning of roles per group provides specificity in our model, ensuring that each group is characterized by distinct roles. It thus lays the foundation for an optimization model that proficiently assigns individuals from $\mathcal{P}$ to specific roles across unique groups, respecting the distinctive role distribution among groups.

4.2 Elaboration on Specific Sets

Expanding on the three principal sets ($\mathcal{P}$, $\mathcal{R}$, and $\mathcal{G}$) introduced in Section 3.1, the optimization model further incorporates several additional, more specific sets to account for individual preferences and group-role associations. These sets serve to refine the model by providing a more personalized approach to role assignment while acknowledging the distinct roles tied to individual groups.

Set $\mathcal{P}_i$, where i is the index of the individual, consists of all individuals that a specific person (i) would prefer to share the same group with. These preferences could stem from various factors, such as compatible skill sets, synergistic work patterns, or pre-existing professional relationships. The model incorporates these preferences to enhance group dynamics and collaboration.

For role preferences, two distinct sets are introduced: $\mathcal{R}_i$ and $\mathcal{R}_i^{\mathcal{N}}$. The set $\mathcal{R}_i$ encompasses all roles that individual i would prefer to be assigned to. These could be roles that align with the individual's skills, interests, or career aspirations. Conversely, the set $\mathcal{R}_i^{\mathcal{N}}$ comprises roles that individual i would prefer not to be assigned to. This could be due to lack of interest, inadequate skill match, or personal aversion towards specific tasks associated with these roles. The model takes both these sets into account when assigning roles to individuals, aiming to maximize satisfaction while maintaining overall efficiency.

A distinct set, $\mathcal{R}_j^{\mathcal{G}}$, with j representing the index of a group, comprises the roles that below to a specific group. Each group must contain at least one role, and each role belongs to only one group. Because of this, the $\mathcal{R}_j^{\mathcal{G}}$ sets–of which there are M–partition the set of roles $\mathcal{R}$.

The nature of the roles encompassed by this set might be uniquely attributed to the requirements of the group it is associated with, defining its identity and functional needs. It is through the medium of these roles, unique to each group, that the model facilitates the assignment of individuals, in a manner that aligns with the aforementioned preferences and aversions.

4.3 Constants and Parameters

This section elaborates on the various constants and parameters that are fundamental to the functioning of the optimization model. These constants and parameters not only establish the constraints of the model but also act as tuning knobs, optimizing the allocation of roles based on individual and group preferences.

The constants max_j and min_j represent the maximum and minimum sizes of each group, respectively. These constants are established by adding together the maximum (max_r) and minimum (min_r) number of people that can be assigned to each role in a given group. Thus, the group size limits are informed by the sum of the role size limits, reflecting the cumulative capacities of the roles within the group. The one running the optimization provides these parameters to ensure group sizes are compatible with operational constraints and optimal team dynamics.

Conversely, max_r and min_r are the maximum and minimum sizes that a specific role can have.

These constants play a pivotal role in maintaining the balance of role assignments, precluding both over- or under-allocation of individuals to a particular role, thus guaranteeing effective role assignment.

From the individual perspective, the constants max_i and min_i represent the upper and lower limits on the number of groups that an individual can participate in. These parameters are designed to account for individual capacity, preventing over-commitment, while also ensuring an individual's meaningful contribution in at least a minimum number of groups.

Another key parameter, δ_i, reflects the extent to which an individual (i) values the roles assigned to them compared to the composition of the group. Ranging from 0 to 1, a δ_i value of 0 indicates that the individual i cares primarily about the group members, while a value of 1 signifies their exclusive concern about their assigned roles. Values in-between signify varying degrees of preference between the group composition and role assignment. This personalized parameter is instrumental in striking a balance between role and group preferences for each individual during the assignment process.

Collectively, these constants and parameters are designed to tailor the optimization model to individual preferences, group dynamics, and operational constraints, ensuring an effective and optimal assignment of individuals to roles within groups.

4.4 Variables

In the optimization model, several key variables operate in conjunction to formulate the optimal assignment of individuals to roles and groups, taking into account individual preferences and constraints. These variables work in synergy to drive the optimization process, with each of them serving a specific purpose within the model.

The first variable, x_{ij}, is a binary variable linking a person (i) to a group (j). This variable takes a value of 1 if person i is assigned to group j, and 0 otherwise. With N individuals and M groups, there are a total of $N \times M$ instances of the x_{ij} variable, representing all possible individual-group assignments in the model.

The second variable, w_{ir}, is a binary variable connecting a person (i) to a role (r). This variable is set to 1 if person i is assigned to role r, and 0 otherwise. Given N individuals and T roles, there are $N \times T$ instances of the w_{ir} variable, encapsulating every potential individual-role pairing within the model.

The third variable, y_{ikj}, is a binary variable that associates an individual (i) with their preferred colleague (k, where $k \in \mathcal{P}_i$) in a specific group (j). This variable takes a value of 1 if person i and person k are assigned to the same group j, and 0 otherwise. Given N individuals, M groups, and an average size of $\mathcal{P}_i$ over all individuals, the total number of instances of the y_{ikj} variable is $N \times M \times$ (average size of $\mathcal{P}_i$), covering every possible individual-individual-group relationship within the model.

The final variable, z_i, represents the count of preferred individuals (from set $\mathcal{P}_i$) assigned to the same group as person i. As it denotes a count, this variable takes on integer values. There is one z_i for each individual, bringing the total number of instances of z_i to N.

4.5 Objective Function

Value in the objective function is derived from two sources: the first one being how many preferred individuals a person is assigned to the same groups with. This is normalized by dividing by the size of their preference list to limit the effect that having very large or very small preference lists will have on the overall solution to the optimization. It can be written mathematically as:

$$\sum_{i=1}^{N} \frac{z_i}{|\mathcal{P}_i|}$$

The second part of the optimization involves the roles that the individuals are assigned to. It can be expressed symbolically as:

$$\sum_{i=1}^{N} \delta_i \left(\frac{\alpha_i a_i}{|\mathcal{R}_i|} + \frac{\gamma_i c_i}{|\mathcal{R}_i^{\mathcal{N}}|} \right)$$

where a_i and c_i are the numbers of preferred roles that the individuals are in. This is found by taking the sum of the role-individual variable w_{ir} over the set of all roles preferred by that individual, and the same is done for the non-preferred roles as well. This can be seen in the symbolical representation below:

$$\sum_{r \in \mathcal{R}_i} w_{ir} = a_i$$

$$\sum_{r \in \mathcal{R}_i^{\mathcal{N}}} w_{ir} = c_i$$

The alpha and gamma are the amounts of positive and negative value that the objective function receives as a result of an individual being assigned to a role that they prefer or one that they do not prefer, respectively. The first summation which indicates the total preferred individuals and second summation which indicates the total preferred and non-preferred roles are then multiplied by $(1-\delta)$ and δ respectively which then changes the weight that each of these two parts contribute to the overall value of the objective. The full formula of the objective function expressed mathematically is thus:

$$\sum_{i=1}^{N} (1-\delta_i) \frac{z_i}{|\mathcal{P}_i|} + \sum_{i=1}^{N} \delta_i \left(\frac{\alpha_i a_i}{|\mathcal{R}_i|} + \frac{\gamma_i c_i}{|\mathcal{R}_i^{\mathcal{N}}|} \right)$$

4.6 Constraints

Constraints (1) – (4) are used to adhere to the size limits of the roles and groups. For groups, the constraint can be represented mathematically as:

$$\sum_{i=1}^{N} x_{ij} \leq max_j \qquad \forall\ j \in \mathcal{G} \tag{1}$$

$$\sum_{i=1}^{N} x_{ij} \geq min_j \qquad \forall\ j \in \mathcal{G} \tag{2}$$

where max_j and min_j are the maximum and minimum sizes of group j, respectively.

These constraints stipulate that the total number of individuals assigned to a specific group j (represented by the summation of x_{ij} over all individuals i) should be within the defined maximum and minimum group sizes (max_j and min_j, respectively).

As for roles, the constraint is defined as:

$$\sum_{i=1}^{N} w_{ir} \leq max_r \qquad \forall\ r \in \mathcal{R} \tag{3}$$

$$\sum_{i=1}^{N} w_{ir} \geq min_r \qquad \forall\ r \in \mathcal{R} \tag{4}$$

where max_r and min_r denote the maximum and minimum capacities of role r, respectively.

According to these constraints, the total number of individuals allocated to a given role r (represented by the sum of w_{ir} across all individuals i) must not exceed the maximum capacity of that role (max_r) and must be at least the minimum required number of individuals for that role (min_r).

Another essential category of constraints in our optimization model relates to the number of roles assigned to each individual. These constraints guarantee that each individual is allocated a manageable number of roles, thereby avoiding overtasking and ensuring their meaningful contribution to each assigned role.

The variable w_{ir}, which indicates the assignment of individual i to role r, is again instrumental here. This constraint involves the summation over all roles r for each specific individual i.

Symbolically, this constraint can be formulated as:

$$\sum_{r=1}^{T} w_{ir} \leq max_i \qquad \forall\, i \in \mathcal{P} \tag{5}$$

$$\sum_{r=1}^{T} w_{ir} \geq min_i \qquad \forall\, i \in \mathcal{P} \tag{6}$$

where max_i and min_i represent the maximum and minimum number of roles an individual i can be assigned, respectively.

As per these constraints, the total number of roles assigned to an individual i (expressed as the sum of w_{ir} over all roles r) should not exceed the maximum number of roles that individual can manage (max_i) and should not be less than the minimum required number of roles (min_i).

The optimization model includes intervariable constraints that link the binary variables y_{ikj}, representing individual preferences, to the integer variables z_i, representing the total number of preferred individuals grouped with a particular individual. Each individual i is associated with one such constraint, ensuring that the total count of preferred individuals assigned to the same group aligns with the individual's preferences.

The variable y_{ikj} is 1 if person i is grouped with person k (a preferred individual in $\mathcal{P}_i$) in group j, and 0 otherwise. Variable z_i, on the other hand, signifies the total count of preferred individuals from set $\mathcal{P}_i$ that are assigned to the same groups as individual i.

The intervariable constraint can be expressed mathematically as:

$$\sum_{j=1}^{M} \sum_{k \in \mathcal{P}_i} y_{ikj} = z_i \qquad \forall\, i \in \mathcal{P} \tag{7}$$

This constraint implies that the summation of y_{ikj} over all groups j and all preferred individuals k in $\mathcal{P}_i$ for a specific individual i is equal to z_i. In other words, the total count of preferred colleagues that are assigned to the same groups as person i (as indicated by z_i) must equal the sum of all instances where person i is grouped with their preferred individuals (y_{ikj} summed over all j and k in $\mathcal{P}_i$).

Such a constraint ensures the consistency between the binary assignment of individuals to preferred groups (y_{ikj}) and the total count of such assignments (z_i), thereby respecting each individual's group preferences in the optimization model. It serves as a crucial link in maintaining the coherence and functionality of the model.

The optimization model further introduces constraints that bridge the relationship between individual-to-group assignments (x_{ij}), preferred individual-to-group assignments (x_{kj}), and individual-to-preferred-individual-to-group assignments (y_{ikj}). These constraints are defined for each combination of an individual i, a preferred individual $k \in \mathcal{P}_i$, and a group j, thereby ensuring that the grouping of an individual with their preferred colleagues is accurately reflected in the model.

The variable y_{ikj} is 1 if person i and their preferred individual $k \in \mathcal{P}_i$ are both assigned to group j, and 0 otherwise. Variables x_{ij} and x_{kj} denote the assignment of individual i and preferred individual k to group j, respectively.

The mathematical representation of the constraint is:

$$y_{ikj} \leq \frac{1}{2}(x_{ij} + x_{kj}) \qquad \forall\, i \in \mathcal{P},\, j \in \mathcal{G},\, k \in \mathcal{P}_i \tag{8}$$

This constraint postulates that y_{ikj} can only be 1 (meaning person i and their preferred individual k are both assigned to group j) when both x_{ij} and x_{kj} equal 1. If either x_{ij} or x_{kj} is not 1 (indicating that either person i or preferred individual k is not assigned to group j), the sum $(x_{ij} + x_{kj})$ is less than 2, and thus half the sum is less than 1. Since y_{ikj} is a binary variable, anything less than 1 necessitates that y_{ikj} must be 0.

By virtue of this constraint, we ensure that an individual i is grouped with their preferred individual k only when both are indeed assigned to the same group j. This constraint plays a vital role in validating the integrity of the model, verifying the consistency of individual preferences, and fostering desirable group compositions.

The final set of constraints in the optimization model interlinks individual role assignments (w_{ir}) and individual group assignments (x_{ij}). Each combination of a person i and a group j is associated with one such constraint. These constraints ensure that any person assigned to a group must also be assigned to at least one role in that group, reinforcing the coherence and consistency of the model.

For a specific individual i and group j, this constraint can be mathematically represented as:

$$\sum_{r \in \mathcal{R}_j^{\mathcal{G}}} w_{ir} \geq x_{ij} \qquad \forall\ i \in \mathcal{P},\ j \in \mathcal{G} \tag{9}$$

This constraint asserts that the sum of w_{ir} over all roles r in group j (roles in set $\mathcal{R}_j^{\mathcal{G}}$) must be greater than or equal to x_{ij}. If a person i is assigned to a group j ($x_{ij} = 1$), then they must also be assigned to at least one role in that group ($\sum_{r \in \mathcal{R}_j^{\mathcal{G}}} w_{ir} > 0$).

This inequality allows for the possibility of a person being assigned to multiple roles within a single group. However, if the user of the algorithm intends to limit each person to a maximum of one role in each group, the inequality can be changed to an equality, thereby restricting the sum of w_{ir} in the group to 1.

By implementing these constraints, the model is able to preserve the practical and operational requirement that each individual assigned to a group also holds a role within that group. This ensures the practical applicability of the model's outputs and facilitates the successful completion of group tasks and objectives.

5 Applications

5.1 Pilot Scheduling

An important use case for this model is the assignment of pilots to flights. In the context of the model in its current state, each flight can be thought of as a group having maximum and minimum sizes of 2 and consisting of 2 roles: pilot and co-pilot, each having a maximum and minimum size of 1. In the present form, this optimization model does not take into account the various rules and regulations that many airlines have for the mandatory time that each pilot must have between the flights that they are operating, which change for the lengths of flights and the time zones that the destination and starting point are, among other things. However, these regulations can be accounted for by simply adding constraints regarding flights such as start and end time zones, length, and time of day etc. to be satisfied by the optimization.

We created a mobile application for this use case. The user interface for this application works by having the administrator creates the groups (flights), with each having an identical set of roles in it (pilot and co-pilot). The pilots are then prompted to give their preferences for the pilots that they would like to fly with, their preferences for which flights they want to fly (as well as whether they want to be pilot or co-pilot on that specific flight), and whether they care more about the pilots they are flying with or the flights that they are flying.

In this way, the preferences of the pilots can be maximized so that as many pilots as possible are assigned to flights that they want to be on and with pilots that they want to fly with. This can improve job satisfaction for the pilots and potentially improve safety for the airline as pilots are more likely to assigned to work with individuals that they work well with. In addition, the pilots' preferences can change between each time that the algorithm is run, so by implementing a rating system for their co-pilots after each flight, the optimization can adapt to changes in the pilots' preferences over time.

5.2 Classroom

This innovative algorithm significantly improves the efficiency and effectiveness of group assignments in the classroom. By taking into account the personal preferences of students regarding their preferred teammates, roles, and groups, it customizes the assignment process to their individual needs. As a result, this increases overall satisfaction, as students are more likely to enjoy working with their preferred peers on roles and projects they are genuinely interested in.

Moreover, it encourages students to be honest about their preferences because they understand that the algorithm's effectiveness relies on authentic information. This honesty enhances the fairness of the system, as it reduces the incentive for attempting to game the system and ensures that all students are treated equitably. Overall, the algorithm's capacity to optimally assign students to groups and roles, coupled with its potential to increase student satisfaction and honesty, makes it an invaluable tool in the classroom setting.

5.3 Corporate Project Assignment

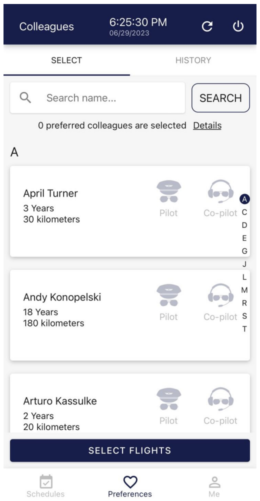
In a corporate setting, this algorithm proves beneficial in assigning individuals to specific roles based on their skills and qualifications. Certain roles require specific expertise; for instance, a project may need a coder, a manager, a graphic designer, and an intern. The algorithm optimizes productivity by favoring the assignment of individuals to the roles they're qualified for, ensuring their skills are used effectively.

The administrator has the ability to modify each individual's non-preferred roles, thereby impacting the overall utility or effectiveness of the assignment. Non-preferred roles in the model are those that, if assigned to an individual, will reduce the assignment's overall utility, manifesting themselves in the real world as decreased productivity and potential dissatisfaction. The model accommodates each individual's unique circumstances, allowing for varying negative impacts from poor role assignments. For example, a coder assigned as a graphic designer might not cause as much disruption as an intern appointed as a manager.

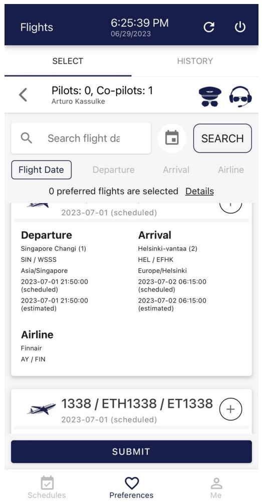
By allowing individuals to choose their preferred roles, coworkers, and projects, the model ensures optimal workplace productivity and efficiency, while also enhancing employee satisfaction. It maintains a balance between business objectives and employee welfare, making it a potent tool for project management in a corporate environment.

6 Pilot Scheduling Data

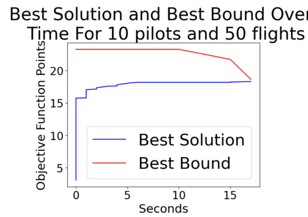
The pilot scheduling problem is an interesting topic for this model because of the possibility of incorporating an airline's further constraints to account for restrictions such as pilots' rest between flights. As a result of this possibility an application has been developed in conjunction with this model in which the pilots can submit their preferences to a server which then runs the optimization and returns the results to the pilots. Images (a) and (b) below show this interface in which a pilot can select their preferences for their colleagues and flights. In addition, the model has primarily been tested on preference data to schedule pilots for flights. The preference data for the simulated pilots is generated randomly based on an arbitrary system of having 4 people with a preference



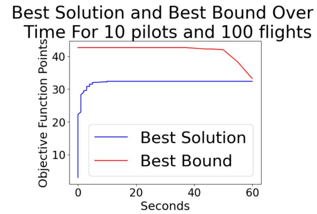
(a) Colleague selection screen



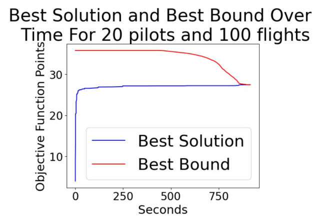
(b) Flight selection screen



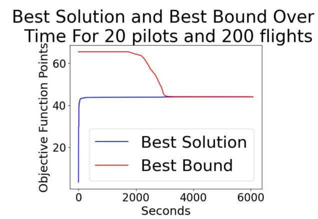
(c) 10 Pilots and 50 Flights



(d) 10 Pilots and 100 Flights



(e) 20 Pilots and 100 Flights



(f) 20 Pilots and 200 Flights

Figure 1: Left: The user interface for the pilot scheduling app
Right: Run time data for the pilot scheduling version of the model on various input sizes

size of 2 and each next 4 people have a preference size of 1 more until the total number of people is reached. For example, for tests with 20 people ($N = 20$) 4 people have a preference size of 2, then 4 have 3, 4 have 4, 4 people have 5 and lastly 4 have 6. This way as the number of people grows, the number of average people in an individual's preference grows as well but much slower, as well as having variation in the size of preference lists. The same system is used for randomizing role preferences ($\mathcal{R}_i$). The δ_i value was set to 0.5 for all individuals so as to make the optimization accommodate all preferences equally. As can be seen from the run time graphs, the model takes a long time to find the optimal solution for large input sizes; however, it reaches an objective value very close to the optimal solution very quickly. This can potentially be exploited to provide a very good close to optimal solution in a fraction of time it takes to find the optimal solution, allowing much larger inputs to be used for which it would be infeasible to try to find the optimal solution.

7 Conclusion

We have presented a optimization model that finds optimal assignments for individuals into roles and groups, considering not only role preferences but also interpersonal preferences. This unique approach transcends traditional methods that primarily focus on role assignments, enhancing both efficiency and satisfaction.

The development of this model began with the fundamental concept of stable matching problems and gradually evolved into an optimization-based approach that better addresses real-world complexities and dynamics. Our approach, more adaptable and comprehensive, is conducive to various scenarios, including but not limited to, business project teams, class groupings, or even complex tasks such as pilot scheduling.

Innovation in this model lies in its flexibility, ability to accommodate individualized preferences, and adaptability to different use cases. By respecting both role suitability and interpersonal affinities, the model potentially leads to higher satisfaction and increased productivity, proving valuable in diverse contexts.

Future work includes extending this model to encompass additional constraints and more complex preferences, as well as finding relaxations to greatly increase the speed of the optimization itself. The continuous integration of human-centric factors into optimization models like ours is indeed a fascinating path to tread, promising significant contributions to both academic research and practical applications. As we move forward, we envision a future where algorithms, such as the one presented here, cater to the nuances of human preferences, yielding optimized results that are harmonious and efficient.

8 Appendix

8.1 Full Formulation

Sets

We have a set $\mathcal{P} = \{1, 2, 3, \ldots, N\}$ of individuals.

We have a set $\mathcal{G} = \{1, 2, 3, \ldots, M\}$ of groups.

We have set $\mathcal{R} = \{1, 2, 3, \ldots, T\}$ of roles.

Specific Sets

Set of person i's preferred team members: $\mathcal{P}_i$, $|\mathcal{P}_i| \leq N$

Set of a person i's preferred roles: $\mathcal{R}_i$, $|\mathcal{R}_i| \leq T$

Set of non-preferred roles: $\mathcal{R}_i^{\mathcal{N}}$, $|\mathcal{P}_i^{\mathcal{N}}| \leq T$

Set of roles in each group: $\mathcal{R}_j^{\mathcal{G}}$, j, $1 \leq |\mathcal{R}_j^{\mathcal{G}}| \leq T - M + 1$

Decision variables:

x_{ij}, binary, $= 1$ if individual i is assigned to group j.

w_{ir}, binary, $= 1$ if individual i is assigned to role r

y_{ikj}, binary, $= 1$ if individual i is assigned to group j, with $k \in \mathcal{P}_i$.

z_i, integer, $= x$ if individual i is assigned to group(s) with x preferred people

a_i, binary, $= 1$ if individual i is assigned to a group j in their preferred groups $\mathcal{G}_i$

c_i, binary, $= 1$ if individual i is assigned to a group j in their non-preferred groups $\mathcal{G}_i^{\mathcal{N}}$

Constants and Parameters

max_i is the maximum number of roles that an individual i can be a part of

min_i is the minimum number of roles that an individual i must be a part of

max_r is the maximum number of individuals that can be in a role r

min_r is the minimum number of individuals that can be in a role r

max_j is the maximum number of individuals that can be in a group j, $max_j := \sum_{r \in \mathcal{R}_j^{\mathcal{G}}} max_r$

min_j is the minimum number of individuals that can be in a group j, $min_j := \sum_{r \in \mathcal{R}_j^{\mathcal{G}}} min_r$

δ_i is each person's parameter preference of team composition vs group: $0 \leq \delta_i \leq 1$

α_i is the benefit in the objective for a person i being assigned to role $r \in \mathcal{R}_i$

γ_i is the loss in the objective for a person i being assigned to a role $r \in \mathcal{R}_i^N$

The overall model:

$$\sum_{i=1}^{N} (1 - \delta_i) \frac{z_i}{|\mathcal{P}_i|} + \sum_{i=1}^{N} \delta_i \left(\frac{\alpha_i a_i}{|\mathcal{R}_i|} + \frac{\gamma_i c_i}{|\mathcal{R}_i^{\mathcal{N}}|} \right) \tag{Objective function}$$

$$s.t. \sum_{j \in \mathcal{R}_i} w_{ir} = a_i \tag{definition of a_i}$$

$$\sum_{j \in \mathcal{R}_i^{\mathcal{N}}} w_{ir} = c_i \tag{definition of c_i}$$

$$\sum_{i=1}^{N} x_{ij} \leq max_j \text{ for each } j \tag{1}$$

$$\sum_{i=1}^{N} x_{ij} \geq min_j \text{ for each } j \tag{2}$$

$$\sum_{r=1}^{T} w_{ir} \leq max_i \text{ for each } i \tag{3}$$

$$\sum_{r=1}^{T} w_{ir} \geq min_i \text{ for each } i \tag{4}$$

$$\sum_{i=1}^{N} w_{ir} \leq max_r \text{ for each } r \tag{5}$$

$$\sum_{i=1}^{N} w_{ir} \geq min_r \text{ for each } r \tag{6}$$

$$\sum_{j=1}^{M} \sum_{k \in \mathcal{P}_i} y_{ikj} = z_i \text{ for each } i \tag{7}$$

$$y_{ikj} \leq \frac{1}{2}(x_{ij} + x_{kj}) \text{ for each } i,\, j,\, k \in \mathcal{P}_i \tag{8}$$

$$\sum_{r \in \mathcal{R}_j^{\mathcal{G}}} w_{ir} \geq x_{ij} \text{ for each } i,\, j \tag{9}$$

References

Cechlárová K, Ferková S (2004) The stable crews problem. *Discrete applied mathematics* 140(1-3):1–17.

Dutta B, Massó J (1997) Stability of matchings when individuals have preferences over colleagues. *journal of economic theory* 75(2):464–475.

Gale D, Shapley LS (1962) College admissions and the stability of marriage. *The American Mathematical Monthly* 69(1):9–15.

Irving RW (1985) An efficient algorithm for the "stable roommates" problem. *Journal of Algorithms* 6(4):577–595.

Lin YT, Huang YM, Cheng SC (2010) An automatic group composition system for composing collaborative learning groups using enhanced particle swarm optimization. *Computers & Education* 55(4):1483–1493.

Rabiei P, Arias-Aranda D, Stantchev V (2023) Introducing a novel multi-objective optimization model for volunteer assignment in the post-disaster phase: Combining fuzzy inference systems with nsga-ii and nrga. *Expert Systems with Applications* 226:120142.

Yekta HA, Bergman D, Day R (2023) Balancing stability and efficiency in team formation as a generalized roommate problem. *Naval Research Logistics (NRL)* 1(70):72–88.