A Variation of the Network Flow Algorithm to Optimize the Diversity of Project Groups

Holger Mauch[1]
[1]*Eckerd College*
mauchh@eckerd.edu

Abstract

Research has shown that the amount of diversity within project groups has an effect on how well groups perform. In particular, groups that are balanced with respect to gender, race, national origin, and social background tend to be more creative, more productive, and perform better on average than non-diversified groups (Hunt 2015, Hunt 2018). Therefore the assignment of employees to project groups must also take into account the sub-goal of diversification during the decision making process (in addition to hard constraints such as suitable skill level of an employee to even qualify for membership in a group.) The additional constraints this imposes on the optimization problem of finding the best group composition can be a considerable challenge for management. In this paper, we present a new variation of a classical network flow algorithm to solve this problem optimally. We show the results of an application area where optimal groups have been formed that work together. The results have been obtained by the software implementation of our algorithm called "NF Group Diversity".

1. Introduction

The motivation for this paper is based on the question "Why do we want diverse groups?" What are the reasons for having diversity in project groups? Or similarly, what are the reasons for having diversity in a classroom? One reason is given in a McKinsey study (Hunt 2015). It found that companies that were in the top 75 percent for ethnic and racial management diversity had profits that were 35 percent higher than those of their industry peers. In addition, companies often can increase their creativity by encouraging diverse opinions and perspectives.

However, managing a diverse group presents numerous challenges for management (Shaban 2016, Saxena 2014). The focus of this paper is to support management with a tool to optimally assign employees from a pool of employees to project groups in such a way that the employees have the skills to work in their assigned group and also to ensure diverse groups at the same time.

To accomplish this we present a new variation of a classical network flow algorithm to match employees from a set E (the employee pool) with open positions in a set of project groups T (teams), such that every employee from E is matched to exactly one project group from T. Every employee has associated with it a skill-preference list which identifies project groups suitable for the employee's skill level. The overall quality of the assignment depends on the extent to which the skill-preferences of the employees are met. Employees also have a set of binary attributes associated with them which hold information about gender, whether they belong to a racial minority, or similar information deemed important for diversity. There is a hard constraint that all project groups should eventually hold the targeted number of employees. Furthermore, there are various additional constraints that promote that project groups contain a mix of employees with certain binary attribute values. We can optimally solve the resulting bipartite matching problem by reducing it to the network flow problem. We successfully designed, implemented and tested the proposed algorithm in a computer software program named "NF Group Diversity".

Please note that the algorithm and software described here helps decision makers to form balanced and diversified groups from a larger pool of individuals (e.g., employees). It is an assignment/matching algorithm, which distributes what we have into heterogeneous groups. If our overall pool of individuals has no or limited diversity to begin with, then there is not much that this algorithm can do for you! There have been several recent studies that suggest strategies on how to obtain a diverse pool of individuals (O'Brien 2015, Hunt 2018), so the reader interested in this type of problem is referred to these.

This paper is organized as follows. We start with a mathematically precise specification of the problem to be solved. Then we provide a description and graphical illustrations of the proposed algorithm that solves the problem. Then we provide one example application area in which the algorithm has been applied successfully. The paper concludes with recommendations and suggested future improvements and variations of the algorithm.

2. Problem Specification

Task: Match employees from a set $E=\{a,b,c,...\}$ (the employee pool) with open positions (seats) $S=\{s_1, s_2, ..., s_n\}$ in a set of project groups $T=\{A,B,...\}$ (teams), such that every employee from E is matched to at most one project group from T.

Every employee e from E has associated with her a skill-preference set $\{c_1, c_2, ..., c_k\}$ which identifies project groups suitable for the employee's skill level. There is the hard constraint that the employee must have the skill set to participate in the assigned team.

Additional constraints reflect the fact that we promote that project groups contain a diverse mix of employees with certain binary attribute values (e.g., male/female). An example target might be to ensure that the female-to-male ratio is approximately the same in each team, thus ensuring diverse teams which reflect the female-to-male ratio of the overall employee pool. The screenshot in Fig. 5 shows such an approach for an employee pool that is roughly 30% male and 70% female. In that example, we specified a "slack", i.e., a maximum deviation from the optimal gender ratio of 0.1=10%. Therefore, a solution to this problem instance (if one exists) will have a male ratio from the interval [30%-10%;30%+10%] = [20%;40%] and the corresponding female ratio from [60%;80%]. The algorithm can easily be adjusted to reflect other diversity goals by changing the color labels of the seat nodes in phase 2 of the algorithm.

The objective of this optimization problem is to maximize the number of employees assigned to teams while respecting all imposed constraints.

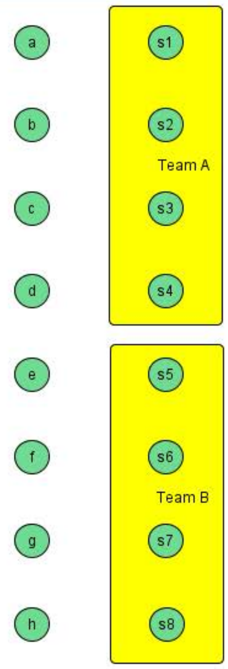

Figure 1. A small Problem Instance

We build a graph model that initially consists of a vertex set that is made up of the set of employees E and the set of seats S as defined above. For illustration purposes an example instance is visualized in Figure 1, where eight employees are to be assigned to two teams of size four each.

3. Description of the Algorithm

The proposed algorithm consists of four phases. There are two novel ideas that are presented in phase 1 and phase 2 of the algorithm. Phases 3 and 4 are an application of existing algorithms that are well-known in the computer science and operations research literature (Cormen et al. 2009, Papadimitriou and Steiglitz 1982).

Phase 1: Recall that every employee has associated with it a skill-preference set $\{c_1, c_2, \ldots, c_k\}$ which identifies project groups suitable for the employee's skill level. We now introduce an edge from an employee to each seat of a team, if the employee has the skills to work in that team.

Phase 2: Characterize (reserve), e.g., by color-coding, each seat in a team as "black", or "white", or "grey" (meaning individuals with either diversity feature can get the seat.) Remove those edges introduced in Phase 1 where the diversity feature value of the employee is incompatible with that of the seat. (Note that by increasing/decreasing the number of "grey" seats we can soften/harden the diversity constraints that we want to achieve. In the example later shown in the screenshot in Fig.5 this is accomplished by adjusting the parameter "Max. Deviation from Gender Ratio".)

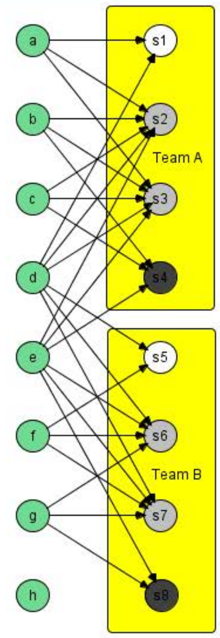

Figure 2. Problem Instance after Phase 2 of the Algorithm

The problem instance shown in Figure 2 shows the resulting graph after phase 2 of the algorithm. For example, employee ***a*** has skills that qualify her for membership in Team ***A*** (but not Team ***B***); moreover, to enforce diversity, employee ***a*** is allowed to occupy the white-colored seat ***s1***, or the grey-colored seats ***s2*** or ***s3***, but not the black-colored seat ***s4***. Therefore, there are three edges originating from vertex ***a*** to the vertices ***s1***, ***s2***, and ***s3***. Note that employee ***h*** does not have the skill set to qualify for membership in either team, thus there are no edges originating from vertex ***h***.

Phase 3: Reduce the resulting Bipartite Matching Problem to the Max-Flow Problem by introducing a source node which connects to all employee vertices ***E*** and a sink node to which all seat vertices ***S*** connect to (Papadimitriou and Steiglitz, 1982). The resulting graph is shown in Figure 3.

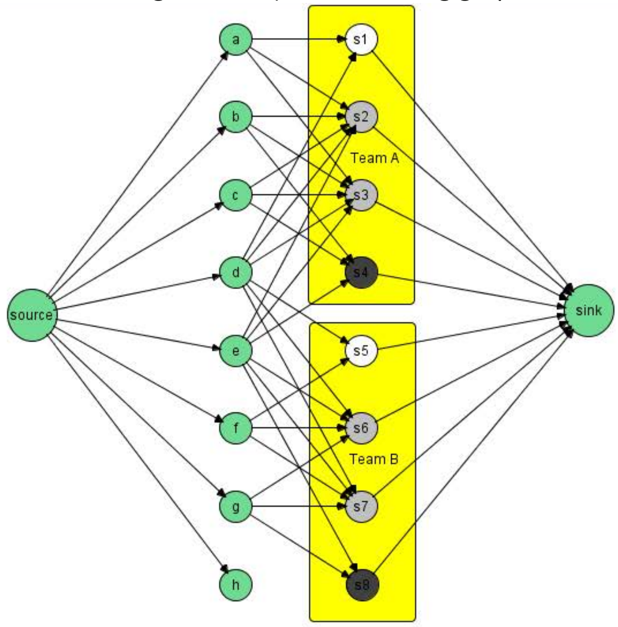

Figure 3. Problem Instance after Phase 3 of the Algorithm

Phase 4: Build an assignment by running a bipartite matching algorithm based on a suitable MaxFlow algorithm: the original algorithm (Ford and Fulkerson, 1962) or an improved version (Edmonds and Karp, 1972), or a more recent variation that solves this well-studied problem.

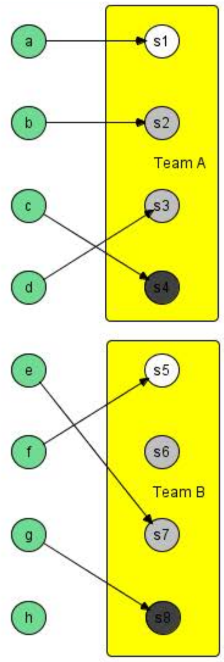

Figure 4. Solved Problem Instance after Phase 4 of the Algorithm

Figure 4 shows the output of phase 4 of the algorithm; for clarity, the source and sink nodes have been removed. In this case, the maximum flow from the source to the sink has a value of 7, and the edges through which this flow takes place (and only these 7 edges) are shown in Figure 4. For example, employee ***a*** has been assigned to seat ***s1*** in team ***A***, employee ***g*** has been assigned to seat ***s8*** in team ***B***, and seat ***s6*** has not been assigned to any employee, because of the lack of a skilled employee to occupy the seat. Note that our algorithm always yields a resulting assignment that is provably optimal (in the sense that the maximum number of seats are assigned) with respect to the constraints imposed. In the small example shown in Figure 3 it is probably impossible to find a flow of value 8. The max flow value is 7.

This algorithm is efficient for all practical numbers of employees and seats. The graph it operates on has $|V|=|E| + |S| + 2$ vertices, and up to $O(ES)$ edges. The time complexity of the algorithm is determined by phase 4 of the algorithm, which takes $O(V(ES)^2)$ if the Edmonds-Karp algorithm is chosen. In our software implementation, the program completes its assignment very quickly, even when the number of employees and seats is in the magnitude of 1000. The partial output of the software for 536 employees and 536 seats is shown in the screenshot of Figure 5 and it has been computed in less than one second on a standard laptop computer.

4. Application: Project Groups

In one practical application of our algorithm, we had an existing employee pool of $|E|$=99 members that had some diversity in several dimensions. Some employees were very skilled and could join a large number of project groups, whereas others were less skilled and could join only a small number of project groups (Table 1).

Table 1. Data of the Employee Pool

Employee ID	Sex	Culture	Skill 1	Skill 2	...	Skill k
a	M	US	A	B	...	K
b	F	US	B	C	...	L
c	M	Int	A	D		
...						

Each team had a limited number of positions (11 seats), and because of budget constraints there was a strict company policy that teams cannot have excess team members. In order to obtain diversity in all teams, we imposed a gender balance constraint and a cultural background (nationality) constraint for most of the seats in a team. Two seats in each team were colored "grey" which softened the diversity constraints imposed on this problem so that all seats could be filled with employees. It is worth to note that with 0% of "grey" seats an optimal matching left some seats unoccupied, which was undesirable for management, so we softened the diversity constraints by repeatedly increasing the number of "grey" seats and re-running the algorithm until we arrived at a solution where all seats were occupied. In this application employees were assigned to at most one team. In the end we achieved pretty well-balanced teams (Table 2) where the female:male ratio was between 18:82 and 37:63 for all teams and coincidentally similar results were obtained for the make-up of the cultural background of each team, where each team roughly had a similar cultural background ratio as the employee pool as a whole.

A theoretical optimum for our employee pool of 99 employees with 27 (27.27%) female members and 31 (31.31%) non-US members that need to be assigned to $|T|=9$ teams of 11 employees each would look as follows: each team would have 3.0 female members, and 3.44 non-US members. This theoretical optimum has not been achieved because of the nature of the input data in terms of the distribution of employee skill level and also because the numbers of team members must always be integers, but with the aforementioned slack we have achieved these results with 2 "grey" seats: Each of the 9 teams had at least 2 and at most 4 female members, and at least 2 and at most 4 non-US members.

In summary, for this application, the algorithm successfully computed an automated assignment of employees to teams, where each employee had the skills needed to work in the assigned team, and there were no longer "all U.S. male" teams as had been the case in the past with a traditional (manual) approach to assigning teams. Of particular practical value turned out to be the fact that for each minority group (female, non-US) the algorithm could ensure at least two members per team from that minority, which avoided the "outsider" feeling that might have arisen if there would have been only a single member from a particular minority.

Table 2. Output of the Algorithm (shown in last column)

Employee	Sex	Culture	Skill 1	Skill 2	...	Skill k	**Assigned Team**
a	M	US	A	B	...	K	**A**
b	F	US	B	C	...	L	**C**
c	M	Int (non-US)	A	D			**D**
...							

5. Conclusions and Future Work

The algorithm introduced in this paper and the software "NF Group Diversity" that implements it (see screenshot in Figure 5) provide an efficient way to support decision makers in finding the best team composition with respect to employee skills and employee diversity. The algorithm should be understood as a building block, which can – often easily - be customized to reflect alternative company policies and diversity goals, e.g., by simply repainting the seats.

Data of the employee pool (such as in Table 1) is delivered to the software in the form of an input file (called input.txt in the screenshot of Figure 5), and then a comprehensive output file with content such as in Table 2 (called output.txt in the screenshot of Figure 5, enriched with additional statistical information) is produced by NF Group Diversity. Improvements of this basic GUI are currently a work in progress.

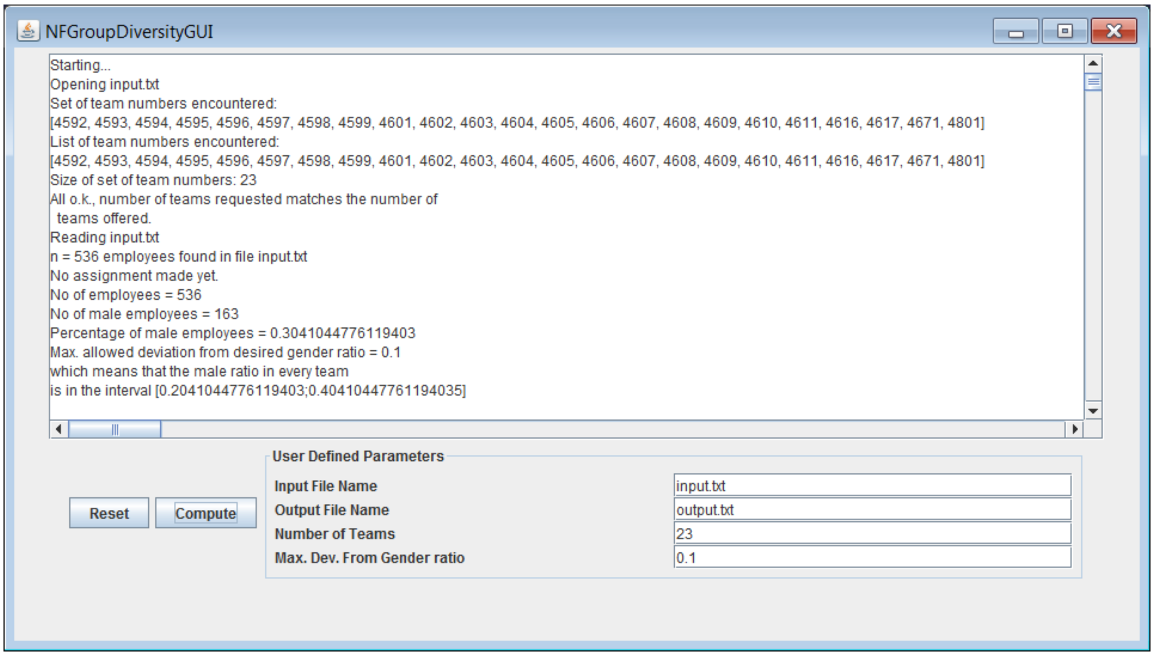

Figure 5. Screenshot of software "NF Group Diversity"

More future work is needed to generalize the algorithm and the software NF Group Diversity so that it can handle non-binary diversity features (say age), and to improve how the seats are colored if there are multiple diversity features to be considered simultaneously. Finally, further efficiency improvements are possible by choosing a more efficient network-flow algorithm in phase 4 of our algorithm.

6. References

Cormen, Thomas H., et al., Introduction to Algorithms, 3rd ed. MIT Press, 2009.

Edmonds, Jack and Richard M. Karp, Theoretical improvements in algorithmic efficiency for network flow problems. Journal of the ACM. 19 (2): 248–264, April 1972.

Ford, L.R. and D.R. Fulkerson, Flows in Networks. Princeton University Press, 1962.

Hunt, Vivian; Layton, Dennis; Prince, Sara, Why diversity matters, McKinsey & Company, January 2015. https://www.mckinsey.com/business-functions/organization/our-insights/why-diversity-matters

Hunt, Vivian, et al., *Delivering through diversity*, McKinsey & Company, January 2018. https://www.mckinsey.com/business-functions/organization/our-insights/delivering-through-diversity

O'Brien, Katherine R., et al., *How to Break the Cycle of Low Workforce Diversity: A Model for Change*, PLoS ONE 10(7): e0133208. doi:10.1371/journal.pone.0133208, July 2015.

Papadimitriou, Christos H. and Kenneth Steiglitz, *Combinatorial Optimization - Algorithms and Complexity*, Prentice Hall, 1982.

Saxena, Sankita, *Workforce Diversity: A Key to Improve Productivity*, Procedia Economics and Finance 11 (2014) 76–85.

Shaban, Aya, *Managing and Leading a Diverse Workforce: One of the Main Challenges in Management*, Procedia Social and Behavioral Sciences 230 (2016) 76–84.