Humboldt-Universität zu Berlin
Mathematisch-Naturwissenschaftliche Fakultät
Institut für Informatik

An interactive system for finding stable student-seminar matchings

Bachelorarbeit

zur Erlangung des akademischen Grades
Bachelor of Science (B. Sc.)

eingereicht von: Aaron Oertel
geboren am: 18.02.1997
geboren in: Dinslaken

Gutachter/innen: Dr. Alexander van der Grinten
Prof. Dr. Henning Meyerhenke

eingereicht am: .... verteidigt am: ....

Contents

1 Abstract

Work in progress

2 Introduction

2.1 Motivation

Many universities require students to enroll in a seminar in order to obtain their degree. Usually the students have a choice between a handful of different seminars, however there are capacity constraints that make it hard to give all students their first choice. Let's consider the following example: 100 students have to be assigned to one of 6 seminars, where each of the seminars has a capacity of 20. The students express their preferences by supplying a strict, but incomplete preference list of the 6 seminars. The goal for the school's administration now is to assign as many students as possible to a seminar of their choice.
What makes this problem harder is that the students preferences aren't necessarily equally distributed. Oftentimes a majority of students prefers one seminar in which case conflicts exist in choosing which students get their first choice. At the same time, it can happen that students go unmatched if their preference lists are short and full of seminars that have reached full capacity already.
The goal of this thesis is to formally model the aforementioned problem, while presenting different algorithms for finding possible matchings and finally to evaluate these algorithms using certain metrics that make sense in the domain of matching, such as stability, rank-maximality, popularity and pareto-optimality. Additionally, an interactive system will be developed, which allows a school's administration to find a student-seminar matching using one of the presented algorithms.

2.2 Formal Definition

The problem of assigning students to seminars can be described as a many-to-one matching, where one seminar is matched to a maximum of c students, where c is the seminar's capacity. Additionally, there are one-sided preferences on the side of the students, who provide a potentially incomplete strict ordering over the seminars. Therefore the problem can be formalized as a bipartite one-to-many matching problem, with one-sided preferences.

2.3 Similar Problems

There are many similar problems, which differ in what the preference lists look like and how many entities of the first set are matched to how many entities of the second set. Studying of these problems reveals some important insights into matching markets.

2.3.1 Stable Marriage Problem

The stable marriage problem was one of the first matching problems to be researched[1] and consequently motivated a lot more research in the domain of matching.
The problem is stated as follows: A set of m men and n women shall be matched

one-to-one, where each men and women provide a complete strict-preference order over the agents of the other set. The deferred acceptance algorithm presented in Gale and Shapley's paper[1] finds a stable, complete matching in polynomial time. Stability is defined as follows: given a women w and any man that she was not matched to m, w does not prefer m more than her current partner, and m does not prefer w more than his current partner.
TODO: present algorithm here?
The algorithm can be executed in two ways:

1. the men have priority, by proposing to women, where the woman has to accept the proposal iff it improves her situation.

2. the women have priority and propose to men. This case is analogue to the first one

It has been shown that all possible executions of the algorithm with men as proposers yield the same stable matching. That matching is men-optimal, which means that every man has the best partner that he can have in any stable matching.[2] Additionally, with men proposing the produced matchings has also been shown to be women-pessimal, meaning that every woman is matched to the worst partner that she can have in any matching.[2]
While the scenario of matching women to men like described (hopefully) doesn't occur in the real world, Gale and Shapley's paper inspired a lot more approaches to other practical problems:

2.3.2 The Hospitals/Residents Problem

A few years after the publication of Gale and Shapley's original paper, it was found that the deferred acceptance algorithm was essentially the same algorithm used by the National Resident Matching Program (NRMP) in the United States to match graduating medical students to residency positions in hospitals.[2] As a matter of fact, Gale and Shapley's paper described an algorithm for the so-called "college admissions problem"[1], which is essentially the same problem. Just like in the stable-marriage problem, there is a solution that be hospital-optimal or resident-pessimal.
The problem can be described as finding a one-to-many matching with two-sided incomplete, but strict preferences. Essentially, a hospital can offer multiple spots and both parties can mark entities from the other set as unacceptable by not including them in their preference list.[3]
In reality, the problem is a bit more complex, as it permits couples of residents to submit preferences together. It has been shown that a stable solution does not always exist and that finding one if it exists or showing that it doesn't exist is NP-complete.[4]
The revised algorithm used by the NRMP uses findings about stability and simple matching markets to find a good approximation, while minimizing opportunities for strategic manipulation, which was indeed possible before.[5]

2.3.3 House allocation problem

2.3.4 Assignment Problem

The problem of matching students to seminars can also be defined as an assignment problem. The goal of the assignment problem is to find a minimum weight perfect matching in a bipartite graph. In this case the goal of the problem is, given the set of students A, seminars B, and a cost function $W : A \times B \rightarrow IR$, to find a map $M : A \rightarrow B$, which minimzes the following objective function: $\sum_{a \in A} W(a, M(a))$.
One of the first algorithms used for solving this problem was the Hungarian algorithm by Munkres, which finds the minimum-weight matching in polynomial time.[6]
Alternatively the problem can be transformed into an instance of the
TODO: add reference to algorithms part.

2.4 Outline

Now that the problem has been formalized and similar problems have been presented, I will present several different algorithmic approaches to find possible matchings. To evaluate these matchings I will also present commonly used metrics like:

- Stability (modified for one sided preferences)
- Rank-maximality (maximum number of students ranked to their first priority)
- Maximum cardinality
- Pareto optimality
- Popularity

Using these metrics I will choose one algorithm for implementation and analyze it's space and time complexities as well as evaluating it's performance using the aforementioned metrics against the other approaches. Next, I will describe the interactive web system developed for using the algorithm and lastly extend the problem to a two-sided problem.

3 Basic Concepts

3.1 Graph Matching

3.1.1 Unweighted Graphs

3.1.2 Weighted Graphs

Let $G = (V, E)$ be an arbitrary edge-weighted graph, so that every edge $e \in E$ has an associated weight $w(e) \in \mathbb{N}$. The weight of a Matching M is now defined as $w(M) = \sum_{e \in M} w(e)$.
In the previous section we have already defined the maximum cardinality matching problem. Using the definition of a Matching's weight we can now define the maximum weight matching problem. The goal of this problem is to find a maximum cardinality matching that maximizes the sum of edge weights.

3.2 Stable Matching

3.2.1 Practical Applications

3.2.2 One to one: Stable Marriage Problem

3.2.3 One to many: Hospital Residents Problem

4 Methodology

test

5 Implementation

test

6 Evaluation

test

7 Discussion

test

8 Conclusion

References

[1] D. Gale and L. S. Shapley, "College admissions and the stability of marriage," *The American Mathematical Monthly*, vol. 69, no. 1, pp. 9–15, 1962.

[2] D. Gusfield and R. W. Irving, *The Stable Marriage Problem: Structure and Algorithms*. Cambridge, MA, USA: MIT Press, 1989.

[3] A. Roth, "The theory and practice of market design," Nobel Prize in Economics documents 2012-5, Nobel Prize Committee, 2012.

[4] E. Ronn, "Np-complete stable matching problems," *Journal of Algorithms*, vol. 11, no. 2, pp. 285 – 304, 1990.

[5] A. E. Roth and E. Peranson, "The redesign of the matching market for american physicians: Some engineering aspects of economic design," Working Paper 6963, National Bureau of Economic Research, February 1999.

[6] J. Munkres, "Algorithms for the assignment and transportation problems," *Journal of the Society for Industrial and Applied Mathematics*, vol. 5, no. 1, pp. 32–38, 1957.

Selbständigkeitserklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbständig verfasst und noch nicht für andere Prüfungen eingereicht habe. Sämtliche Quellen einschließlich Internetquellen, die unverändert oder abgewandelt wiedergegeben werden, insbesondere Quellen für Texte, Grafiken, Tabellen und Bilder, sind als solche kenntlich gemacht. Mir ist bekannt, dass bei Verstößen gegen diese Grundsätze ein Verfahren wegen Täuschungsversuchs bzw. Täuschung eingeleitet wird.

Berlin, den May 11, 2019 ....