Integer programming formulations for college admissions with ties

Kolos Csaba Ágoston[1,2], Péter Biró[1,2], Endre Kováts[3], and Zsuzsanna Jankó[4]

[1] Institute of Economics, Hungarian Academy of Sciences, Hungary
[2] Department of Operations Research and Actuarial Sciences, Corvinus University of Budapest, Hungary
[3] Budapest University of Technology and Economics
[4] University of Hamburg, Germany
peter.biro@krtk.mta.hu

Abstract. When two students with the same score are competing for the last slot at a university programme in a central admission scheme then different policies may apply across countries. In Ireland only one of these students is admitted by a lottery. In Chile both students are admitted by slightly violating the quota of the programme. Finally, in Hungary none of them is admitted, leaving one slot empty. We describe the solution by the Hungarian policy with various integer programing formulations and test them on a real data from 2008 with around 100,000 students. The simulations show that the usage of binary cutoff-score variables is the most efficient way to solve this problem when using IP technique. We also compare the solutions obtained on this problem instance by different admission policies. Although these solutions are possible to compute efficiently with simpler methods based on the Gale-Shapley algorithm, our result becomes relevant when additional constraints are implied or more complex goals are aimed, as it happens in Hungary where at least three other special features are present: lower quotas for the programmes, common quotas and paired applications for teachers studies.

Keywords: integer programming · college admissions · stable matching.

1 Introduction

Gale and Shapley gave a standard model for college admissions [15], where stable matching is was the solution concept suggested. Intuitively speaking a matching is stable if the rejection of an application at a college is explained by the saturation of that college with higher ranked students. Gale and Shapley showed that a stable matching can always be found by their so-called deferred-acceptance algorithm, which runs in linear time in the number of applications, see e.g. [16]. Moreover, the student-oriented variant results in the student-optimal stable matching, which means that no student could get a better assignment in any other stable matching. The theory of stable matchings have been intensively studied since 1962 by mathematicians/computer scientists (see e.g. [16])

and economists/game theorists (see e.g. [20]). The Gale-Shapley algorithm has also been used in practice all around the world [8], first in 1952 in the US resident allocation programme, called NRMP [18], then also in school choice, e.g. in Boston [1] and New York [2]. In Hungary, the national admission scheme for secondary schools follows the original Gale-Shapley model and algorithm [9], and the higher education admission scheme also uses a heuristic based on the Gale-Shapley algorithm [10].

The Hungarian higher education admission scheme have at least four important special features: the presence of ties, the lower and common quotas, and the paired applications. Each of the latter three special features makes the problem NP-hard [11], only the case of ties is resolvable efficiently [12]. In a recent paper [4] we studied the usage of integer programming techniques for finding stable solutions with regard to each of these four special features separately, and we managed to solve the case of lower quotas for the real instance of 2008. In this follow-up work we develop and test new IP formulations for the case of ties. The ultimate goal of this line of work is to suggest a solution concept for the college admission problem where ties and common quotas are also present, together with providing integer programming formulations that are suitable to compute this solution for large scale applications, such as the Hungarian university admission scheme with over 100,000 students.

First we start by investigating the basic Gale-Shapley model and then we consider the case of ties. **Due to the space limit we defer the description of IPs to the full version of the paper, here we present only the results of the simulations.**

2 Model descriptions

In this section first we present the classical Gale-Shapley college admission problem and then the case of ties.

2.1 The Gale-Shapley model

In the classical college admissions problem by Gale and Shapley [15] the students are matched to colleges.[5] In our paper we will refer the two sets as *applicants* $A = \{a_1, \ldots, a_n\}$ and *colleges* $C = \{c_1, \ldots c_m\}$. Let u_j denote the upper quota of college c_j. Regarding the preferences, we assume that the applicants provide strict rankings over the colleges, where r_{ij} denotes the ranking of the application (a_i, c_j) in applicant a_i's preference list. We suppose that the students are ranked according to their scores at the colleges, so college c_j ranks applicant a_i according to her score s_{ij}, where higher score is better. Let $E \subseteq A \times C$ denote the set of applications. A *matching* is a set of applications, where each student is admitted to at most one college and each college has at most as many assignees

[5] In the computer science literature this problem setting is typically called Hospital / Residents problem (HR), due to the National Resident Matching Program (NRMP) and other related applications.

and its quota, u_j. For a matching M let $M(a_i)$ denote the college where a_i is admitted (or $\emptyset$ if a_i is not allocated to any college) and let $M(c_j)$ denote the set of applicants admitted to c_j in M. A matching $M \subset E$ is *stable* if for any application (a_i, c_j) outside M either a_i prefers $M(a_i)$ to c_j or c_j filled its seats with u_j applicants who all have higher scores than a_i has. The deferred-acceptance algorithm of Gale and Shapley provides a student-optimal stable matching in linear time [15].

The notion of *cutoff scores* is important for both the classical Gale-Shapley model and its generalisations with ties and common quotas. Let t_j denote the cutoff score of college c_j and let $\mathbf{t}$ denote a set of cutoff scores. We say that matching M is implied by cutoff scores $\mathbf{t}$ if every student is admitted to the most preferred college in her list, where she achieved the cutoff score. We say that a set of cutoff scores $\mathbf{t}$ *corresponds* to a matching M if $\mathbf{t}$ implies M. For a matching M an applicant a_i has *justified envy* towards another applicant a_k at college c_j if $M(a_k) = c_j$, a_i prefers c_j to $M(a_i)$ and a_i is ranked higher than a_k at c_j (i.e. $s_{ij} > s_{kj}$). A matching with no justified envy is called *envy-free* (see [22] and [21]).

It is not hard to see that a matching is envy-free if and only if it is implied by some cutoff scores [3]. Note that an envy-free matching might not be stable because of blocking with empty seats, i.e. when a student a_i prefers c_j to $M(a_i)$ and c_j is not saturated (i.e. $|M(c_j)| < u_j$). In this case a matching is called *wasteful*. Again, by definition it follows that a matching is stable if and only if it is envy-free and non-wasteful (see also [6]). To achieve non-wastefulness we can require the cutoff of any unsaturated college to be minimum (zero in our case). Alternatively we may require that no cutoff score may be decreased without violating the quota of that college, while keeping the other cutoff scores. Furthermore, we may also satisfy the latter condition by ensuring that we select the student-optimal envy-free matching, which is the same as the student-optimal stable matching [22]. To return this solution we only need to use an appropriate objective function. We will use the above described connections when developing our IPs.

2.2 Case of ties

In many nationwide college admission programmes the students are ranked based on their scores, and ties may appear. In Hungary, for instance, the students can obtain integer points between 0 and 500 (the maximum was 144 until 2007), so ties do occur. When ties are present then one way to resolve this issue is to break ties by lotteries, as done in Ireland (so a lucky student with 480 point may be admitted to law studies, whilst an unlucky student with the same score may be rejected). However, the usage of lotteries can be seen unfair, so in some countries, such as Hungary [12] and Chile [17] equal treatment policies are used, meaning that students with the same score are either all accepted or all rejected. In case of such a policy, there are two reasonable variants when deciding about the last group of students without whom the quota is unfilled and with whom the quota is violated. In the restrictive policy, used in Hungary, the quotas are never violated, so this last group of students is always rejected, whilst in Chile they use a permissive policy and they always admit this last group of students. For instance, if there are three students, a_1, a_2 and a_3, applying to a programme of quota 2 with scores 450, 443, and 443, respectively then in Hungary only a_1 is admitted, whilst in Chile all three students are admitted. In Ireland, a_1 is admitted and they use a lottery to decide whether a_2 or a_3 will get the last seat.

Stable matchings for the case of ties were defined through the cutoff scores in [12]. The usage of cutoff scores in case of ties make the solution *envy-free*, meaning that no student a_i may be rejected from college c_j if this college admitted another student with score equal to or lower than the score of student a_i. This allocation concept is called also equal treatment policy, as the admission of a student to a programme implies the admission offer to all other students with the same score. A matching is envy-free for college admission problem with ties if and only if it is induced by cutoff scores [3]

For the restrictive policy used in Hungary, the stability of the matching can be defined by adding a non-wastefulness condition to envy-freeness. Namely, a matching induced by cutoff scores is *stable* if no college can decrease its cutoff score without violating its quota, assuming that the other cutoff scores remain the same. In the more permissive Chilean policy a matching is stable if by decreasing the cutoff score of any college there would be empty seats left there. (We note that the stability of a matching can be equivalently defined by the lack of a set of blocking applications involving one college and a set of applicants such that this set of applications would be accepted by all parties when compared to the applications of the matching considered. See more about this connection in [14].)

Biró and Kiselgof [12] showed two main theorems about stable matchings for college admissions with ties. In their first theorem they showed that a student-optimal and a student-pessimal stable matchings exist for both the restrictive policy (Hungary) and the permissive policy (Chile), where the cutoff scores are minimal / maximal, respectively. Furthermore, they also proved the intuitive results that if we compare the student-optimal cutoff scores (or the student-pessimal ones) with respect to the three reasonable policies, namely the Hungarian (restrictive), the Irish (lottery), and the Chilean (permissive), then the Hungarian cutoff scores are always as high for each college than the Chilean cutoff scores and the Irish cutoff scores are in between these. When considering the student-optimal stable matching, it turns out to be also the student-optimal envy-free matching, as described in [3].

3 Simulations

In this section we present the main simulation results.

3.1 Gale-Shapley model

We took the 2008 data after breaking the ties randomly, by considering only the faculty quotas and keeping only the highest ranked application of each student for every programme (i.e. the application for either the state funded or the privately funded seat). We used AMPL with Gurobi for solving the IPs.

IP formulations	#variables	#constraints	#non-0 elem.	size(Kb)	run time(s)
SO-BB	287,035	381,115	73,989,595	1,319,663	1,139
SO-NW-CUT	291,935	673,050	2,463,808	69,464	81
MIN-CUT	289,485	668,150	2,169,423	64,254	5,062
MSMR-CUT	289,485	668,150	2,169,423	69,846	2,318
SO-NW-BIN-CUT	574,070	955,185	3,028,078	75,810	107
MIN-BIN-CUT	574,070	952,735	2,738,593	65,657	871
MSMR-BIN-CUT	574,070	952,735	2,738,593	66,467	4,325
MSMR-EF	n.a.	n.a.	n.a.	8,667,403	n.a.

3.2 Case of ties

We used the 2008 data with the original ties by considering again the faculty quotas and keeping only the highest ranked application of each student for every programme.

IP formulations	#variables	#constraints	#non-0 elem.	size(Mb)	run time(s)
MIN-CUT	289,485	668,150	2,169,423	59,694	5,247
MSMR-CUT	289,485	668,150	2,169,423	65,286	1,460
MIN-BIN-CUT	428,513	807,178	2,447,479	53,548	982
MSMR-BIN-CUT	428,513	807,178	2,447,479	57,106	1,362
SO-H-NW-CUT	578,970	1,694,333	4,793,409	114,882	1,310
SO-H-NW-BIN-CUT	861,105	1,813,840	5,352,772	118,828	165

Finally, we conducted the simulation on the same 2008 data, where we compared the results for the Hungarian, Irish and Chilean policies. The results indeed follow the theorems of [12] regarding the cutoff scores for the three different policies. The most interesting fact of the simulation is that for the Hungarian and Irish policies the difference between the student-optimal and student-pessimal solutions is minor, as demonstrated also in earlier paper for large markets, such as [19]. However, for the Chilean policy this difference was more significant.

	size of matching		average rank		average cutoffs		# rejections	
policies	A-opt.	C-opt.	A-opt.	C-opt.	A-opt.	C-opt.	A-opt.	C-opt.
Hungarian	86,195	86,195	1.2979	1.2979	58.3931	58.3931	37,698	37,698
Irish	86,410	86,410	1.2916	1.2916	58.2090	58.2106	36,802	36,804
Chilean	86,650	86,614	1.2824	1.2844	57.2502	57.5200	35,668	35,901

Acknowledgements

The authors were supported by the Hungarian Academy of Sciences under its Momentum Programme (LP2016-3/2018) and Cooperation of Excellences Grant (KEP-6/2018).

References

1. Abdulkadiroğlu A., Pathak A., and Roth A.E. The New York City high school match. *American Economic Review, P&P*, 95(2):364–367, 2005.
2. Abdulkadiroğlu A., Pathak P.A., Roth A.E., and Sönmez T. The Boston public school match. *American Economic Review, P&P*, 95(2):368–371, 2005.
3. Ágoston K.Cs., and Biró P. Modelling Preference Ties and Equal Treatment Policy. *In Proceedings of ECMS 2017: 31st European Conference on Modelling and Simulation*, pages 516-522, 2017.
4. Ágoston K.Cs., Biró P., and McBride I. Integer programming methods for special college admissions problems. *Journal of Combinatorial Optimization*, 32(4):1371–1399, 2016.
5. Ágoston K.Cs., Biró P., and Szántó R. Stable project allocation under distributional constraints. *Operations Research Perspectives*, 5:59–68, 2018.
6. Azevedo, E.M., and Leshno, J.D. A supply and demand framework for two-sided matching markets. *Journal of Political Economy*, 124(5):1235–1268, 2016.
7. Baïou M., and Balinski M. The stable admissions polytope. *Mathematical Programming*, 87(3), Ser. A:427–439, 2000.
8. Biró, P. Applications of matching models under preferences. In Endriss, U., editor, *Trends in Computational Social Choice*, chapter 18, pages 345–373. AI Access. 2017.
9. Biró P. Matching Practices for Secondary Schools – Hungary. matching-in-practice.eu, accessed on 23 August 2014.
10. Biró P. University admission practices - Hungary. matching-in-practice.eu, accessed on 23 August 2014.
11. Biró, P., Fleiner, T., Irving, R.W, and Manlove, D.F. The College admissions problem with lower and common quotas. *Theoretical Computer Science*, 411:3136–3153, 2010.
12. Biró, P., and Kiselgof, S. College admissions with stable score-limit. *Central European Journal of Operations Research*, 23(4):727–741, 2015.
13. Biró P., McBride I., and Manlove D.F. The Hospitals / Residents problem with Couples: Complexity and Integer Programming models. *In Proceedings of SEA 2014: the 13th International Symposium on Experimental Algorithms*, vol 8504 of LNCS, pp 10-21, Springer, 2014.
14. Fleiner T., and Jankó Zs. Choice Function-Based Two-Sided Markets: Stability, Lattice Property, Path Independence and Algorithms. *Algorithms*, 7(1):32–59, 2014.
15. Gale D., and Shapley L.S. College Admissions and the Stability of Marriage. *American Mathematical Monthly*, 69(1):9–15, 1962.
16. Manlove D.F. Algorithms of Matching Under Preferences. *World Scientific Publishing*, 2013.
17. Rios I., Larroucau T., Parra G., and Cominetti R. College admissions problem with ties and flexible quotas. *Working paper*, 2014.
18. Roth A.E. The evolution of the labor market for medical interns and residents: a case study in game theory. *Journal of Political Economy*, 6(4):991–1016, 1984.
19. Roth A.E., and Peranson E. The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design. *American Economic Review*, 89:748–780, 1999.
20. Roth A.E., and Sotomayor M.A.O. Two-sided matching: a study in game-theoretic modeling and analysis. *Cambridge: Econometric Society monographs*, 1990.
21. Yokoi Y. Envy-Free Matchings with Lower Quotas. arXiv preprint, arXiv:1704.04888, 2017.
22. Wu Q., and Roth A.E. The lattice of envy-free matchings. mimeo, 2016.