Journal of Discrete Algorithms 13 (2012) 59–66

Contents lists available at SciVerse ScienceDirect

Journal of Discrete Algorithms

www.elsevier.com/locate/jda

Improved approximation bounds for the Student-Project Allocation problem with preferences over projects ☆

Kazuo Iwama [a], Shuichi Miyazaki [b,*], Hiroki Yanagisawa [c]

[a] Graduate School of Informatics, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto-shi, Kyoto 606-8501, Japan
[b] Academic Center for Computing and Media Studies, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto-shi, Kyoto 606-8501, Japan
[c] IBM Research – Tokyo, 1623-14 Shimotsuruma, Yamato-shi, Kanagawa 242-8502, Japan

ARTICLE INFO

Article history:
Received 24 June 2011
Received in revised form 6 January 2012
Accepted 20 February 2012
Available online 23 February 2012

Keywords:
The student project allocation problem
Stable matching
NP-hardness
Approximation algorithm
Approximation ratio

ABSTRACT

Manlove and O'Malley (2008) [8] proposed the Student-Project Allocation problem with Preferences over Projects (SPA-P). They proved that the problem of finding a maximum stable matching in SPA-P is APX-hard and gave a polynomial-time 2-approximation algorithm. In this paper, we give an improved upper bound of 1.5 and a lower bound of 21/19 (> 1.1052).

1. Introduction

Assignment problems based on the preferences of participants, which originated from the famous *Hospitals/Residents problem* (*HR*) [3], are important almost everywhere, such as in education systems where students must be allocated to elementary schools or university students to projects. In the university case, each student may have preferences over certain research projects supervised by professors and usually there is an upper bound on the number of students each project can accept. Our basic goal is to find a "stable" allocation where no students (or projects or professors if they also have preferences over students) can complain of unfairness. This notion of stability was first introduced by Gale and Shapley in the context of the famous *Stable Marriage problem* in 1962 [2].

The *Student-Project Allocation problem* (*SPA*) is a typical formulation of this kind of problem originally described by Abraham, Irving, and Manlove [1]. The participants here are *students*, *projects*, and *lecturers*. Each project is offered by a single lecturer, though one lecturer may offer multiple projects. Each project and each lecturer has a *capacity*. Students have preferences over projects, and lecturers have preferences over students. Our goal is to find a stable matching between students and projects satisfying all of the capacity constraints for projects and lecturers. They proved that all stable matchings for a single instance have the same size, and proposed linear-time algorithms to find one [1].

Manlove and O'Malley [8] proposed a variant of SPA, called *SPA with Preferences over Projects* (*SPA-P*), where lecturers have preferences over projects they offer rather than preferences over students. In contrast to SPA, they pointed out that the sizes of stable matchings may differ, and proved that the problem of finding a maximum stable matching in SPA-P, denoted

☆ A preliminary version of this paper was presented at the 8th Annual Conference on Theory and Applications of Models of Computation, TAMC 2011. This work was supported by KAKENHI 22240001 and 20700009.

* Corresponding author.
E-mail addresses: iwama@kuis.kyoto-u.ac.jp (K. Iwama), shuichi@media.kyoto-u.ac.jp (S. Miyazaki), yanagis@jp.ibm.com (H. Yanagisawa).

1570-8667/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.jda.2012.02.001

MAX-SPA-P, is APX-hard. They also presented a polynomial-time 2-approximation algorithm. Specifically, they provided a polynomial-time algorithm that finds a stable matching, and proved that any two stable matchings differ in size by at most a factor of two.

1.1. Our contributions

In this paper, we improve both the upper and lower bounds on the approximation ratio for MAX-SPA-P. We give an upper bound of 1.5 and a lower bound of 21/19 (> 1.1052) (under the condition that P ≠ NP). For the upper bound, we modify Manlove and O'Malley's algorithm SPA-P-APPROX [8] using Király's idea [7] for the approximation algorithm to find a maximum stable matching in a variant of the stable marriage problem (*MAX-SMTI*). We also show that our analysis is tight. For the lower bound, we give a gap-preserving reduction from (a variant of) MAX-SMTI. Our reduction also gives a lower bound of 1.25 under the Unique Games Conjecture.

2. Preliminaries

Here we give a formal definition of SPA-P and MAX-SPA-P, derived directly from the literature [8]. An instance I of SPA-P consists of a set S of *students*, a set P of *projects*, and a set L of *lecturers*. Each lecturer $\ell_k \in L$ offers a subset P_k of projects. Each project is offered by exactly one lecturer, i.e., $P_{k_1} \cap P_{k_2} = \emptyset$ if $k_1 \neq k_2$. Each student $s_i \in S$ has an *acceptable* set of projects, denoted A_i, and has a strict order on A_i according to preferences. Each lecturer ℓ_k also has a strict order on P_k according to preferences. Also, each project p_j and each lecturer ℓ_k has a positive integer, called a *capacity*, denoted c_j and d_k, respectively.

An *assignment* M is a subset of $S \times P$ where $(s_i, p_j) \in M$ implies $p_j \in A_i$. Let $(s_i, p_j) \in M$ and ℓ_k be the lecturer who offers p_j. Then we say that s_i *is assigned to* p_j in M, and p_j *is assigned* s_i in M. We also say that s_i *is assigned to* ℓ_k in M and ℓ_k *is assigned* s_i in M.

For $s \in S$, let $M(s)$ be the set of projects to which s is assigned in M. For $r \in P \cup L$, let $M(r)$ be the set of students assigned to r in M. If $M(s_i) = \emptyset$, we say that the student s_i is *unassigned* in M, otherwise s_i is *assigned* in M. We say that the project p_j is *under-subscribed*, *full*, or *over-subscribed* with respect to M according to whether $|M(p_j)| < c_j$, $|M(p_j)| = c_j$, or $|M(p_j)| > c_j$, respectively, under M. If $|M(p_j)| > 0$, we say that p_j is *non-empty*, otherwise, it is *empty*. Corresponding definitions apply to each lecturer ℓ.

A *matching* M is an assignment such that $|M(s_i)| \leqslant 1$ for each s_i, $|M(p_j)| \leqslant c_j$ for each p_j, and $|M(\ell_k)| \leqslant d_k$ for each ℓ_k. For a matching M, if $|M(s_i)| = 1$, we may use $M(s_i)$ to denote the unique project to which s_i is assigned. The *size* of a matching M, denoted $|M|$, is the number of students assigned in M.

Given a matching M, a (student, project) pair (s_i, p_j) *blocks* M, or is a *blocking pair* for M, if the following three conditions are met:

1. $p_j \in A_i$.
2. Either s_i is unassigned or s_i prefers p_j to $M(s_i)$.
3. p_j is under-subscribed and either
   (a) $s_i \in M(\ell_k)$ and ℓ_k prefers p_j to $M(s_i)$, or
   (b) $s_i \notin M(\ell_k)$ and ℓ_k is under-subscribed, or
   (c) $s_i \notin M(\ell_k)$, ℓ_k is full, and ℓ_k prefers p_j to ℓ_k's worst non-empty project,
   where ℓ_k is the lecturer who offers p_j.

Given a matching M, a *coalition* is a set of students $\{s_{i_0}, s_{i_1}, \ldots, s_{i_{r-1}}\}$ for some $r \geqslant 2$ such that each s_{i_j} is assigned in M and prefers $M(s_{i_{j+1}})$ to $M(s_{i_j})$, where $j+1$ is taken modulo r. A matching that has no blocking pair nor coalition is *stable*. Refer to [8] for the validity of this definition of stability. SPA-P is the problem of finding a stable matching, and MAX-SPA-P is the problem of finding a maximum stable matching.

We say that A is an r-approximation algorithm if it satisfies $OPT(I)/A(I) \leqslant r$ for all instances I, where $OPT(I)$ and $A(I)$ are the sizes of the optimal and the algorithm's solutions for I, respectively.

3. Approximability

3.1. Algorithm SPA-P-APPROX-PROMOTION

Manlove and O'Malley's algorithm SPA-P-APPROX [8] proceeds as follows. First, all students are unassigned. Any student (s) who has non-empty preference list applies to the top project (p) on the current list of s. If the lecturer (ℓ) who offers p has no incentive to accept s for p, then s is rejected. When rejected, s deletes p from the list. Otherwise, (s, p) is added to the current matching. If, as a result, ℓ becomes over-subscribed, ℓ rejects a student from ℓ's worst non-empty project to satisfy the capacity constraint. This continues until there is no unassigned student whose preference list is non-empty. Manlove and O'Malley proved that the obtained matching is stable.

We extend SPA-P-APPROX using Király's idea [7]. During the execution of our algorithm SPA-P-APPROX-PROMOTION, each student has one of two states, "unpromoted" or "promoted". At the beginning, all of the students are unpromoted. The application sequence is unchanged. When a student (s) becomes unassigned with her preference list exhausted, s is promoted.