

# Finding Social Optima in Congestion Games with Positive Externalities

Bart de Keijzer and Guido Schäfer

Centrum Wiskunde & Informatica (CWI), Amsterdam  
B.de.Keijzer@cwi.nl, G.Schaefer@cwi.nl

**Abstract.** We consider a variant of congestion games where every player  $i$  expresses for each resource  $e$  and player  $j$  a positive *externality*, i.e., a value for being on  $e$  together with player  $j$ . Rather than adopting a game-theoretic perspective, we take an optimization point of view and consider the problem of optimizing the social welfare.

We show that this problem is NP-hard even for very special cases, notably also for the case where the players' utility functions for each resource are affine (contrasting with the tractable case of linear functions [3]). We derive a 2-approximation algorithm by rounding an optimal solution of a natural LP formulation of the problem. Our rounding procedure is sophisticated because it needs to take care of the dependencies between the players resulting from the pairwise externalities. We also show that this is essentially best possible by showing that the integrality gap of the LP is close to 2.

Small adaptations of our rounding approach enable us to derive approximation algorithms for several generalizations of the problem. Most notably, we obtain an  $(r+1)$ -approximation when every player may express for each resource externalities on player sets of size  $r$ . Further, we derive a 2-approximation when the strategy sets of the players are restricted and a  $\frac{3}{2}$ -approximation when these sets are of size 2.

## 1 Introduction

Congestion games constitute an important class of non-cooperative games that model situations in which  $n$  players compete for the usage of  $m$  resources. Every player  $i$  selects a subset of resources from a collection of resource subsets that are available to him. The utility  $u_{i,e}(x)$  that player  $i$  receives for resource  $e$  depends on the total number  $x$  of players who have chosen this resource. Rosenthal's original definition of congestion games [13] assumes that the utility functions of the players are identical for each resource, i.e.,  $u_{i,e} = u_{j,e}$  for every two players  $i, j$  and every resource  $e$ . Milchtaich [11] introduced the more general congestion game model where every player  $i$  has a *player-specific* utility functions  $u_{i,e}$  as described above.

Ever since their introduction in 1973, congestion games have been the subject of intensive research in game theory and, more recently, in algorithmic game theory. Most of these studies adopt a distributed viewpoint and focus on issues

like the existence and inefficiency of Nash equilibria, the computational complexity of finding such equilibria, etc. (see, e.g., [12] for an overview). Much less attention has been given to the study of congestion games from a centralized viewpoint.

Studying these congestion games from a centralized viewpoint is important in situations where a centralized authority has influence over the players in the game. Also, adopting a centralized perspective may help in acquiring insights about the decentralized setting: if it is hard to find an (approximate) optimum or near-optimum in the centralized case where all the players are completely coordinated, it certainly will be hard for the players to reach such a solution in the decentralized case, where besides lack of coordinated computation, additional issues related to selfishness and stability arise. Lastly, we believe that studying this optimization problem is interesting for its own sake, as it can be seen as a generalization of various fundamental optimization problems.

In this paper, we are interested in the problem of computing an assignment of players to resources such that the *social welfare*, i.e., the sum of the utilities of the players, is maximized. We are aware only of two related articles [3,4] (see paragraph on related work for more details) that study player-specific congestion games from an optimization perspective. Both articles assume that the players are *anonymous* [3] in the sense that the utility function  $u_{i,e}$  of a player  $i$  only depends on the number of players using resource  $e$ , but not on their identities.

The assumption that all players are anonymous is overly simplistic in many situations. We therefore extend the player-specific congestion game model of Milchtaich [11] to incorporate *non-anonymous* players. More specifically, let  $N = [n]$  be the set of players<sup>1</sup> and suppose that player  $i$ 's value for sharing resource  $e$  with player  $j$  is  $v_{ije}$  (possibly  $j = i$ ). We define the utility of player  $i$  for resource  $e$  and player set  $S \subseteq N$  as  $u_{i,e}(S) = \sum_{j \in S} v_{ije}$ . We refer to these games as *generalized congestion games*. The externality  $v_{ije}$  that player  $j$  imposes on player  $i$  on resource  $e$  can be negative or positive. We speak of a generalized congestion game with *positive*, *negative* or *mixed externalities*, respectively, when the  $v_{ije}$ 's are positive, negative or arbitrary.

**Our Contributions.** We study the problem of computing a welfare maximizing assignment for generalized congestion games. As in [3,4], we concentrate on the case where each of the  $n$  players has to choose one of  $m$  available resources (symmetric players, singleton resource sets).

We first consider the problem with mixed externalities and show that it is strongly NP-hard and  $n^{1-\varepsilon}$ -inapproximable for every  $\varepsilon > 0$  even for  $m = 2$  resources. We also give a polynomial-time algorithm that solves the problem when the number of players is constant.

In light of this inapproximability result, we then focus on the problem of computing an optimal assignment for generalized congestion games with positive externalities (MAX-CG-POS-EXT). We derive a polynomial-time algorithm that solves MAX-CG-POS-EXT for  $m = 2$  resources. We show that MAX-CG-POS-EXT

---

<sup>1</sup> We use the notation  $[k]$  to denote the set  $\{1, \dots, k\}$  for a positive integer  $k$ .

is strongly NP-hard for  $m \geq 3$  resources and therefore focus on approximation algorithms.

We derive a deterministic 2-approximation algorithm for MAX-CG-POS-EXT. Our algorithm computes an optimal solution to a natural LP-relaxation of the problem and then iteratively rounds this solution to an integer solution, thereby losing at most a factor 2 in the objective function value. We also show that the integrality gap of the underlying LP is close to 2 and therefore the approximation factor of our algorithm is essentially best possible.

Our rounding procedure is sophisticated because it needs to take care of the dependencies between the players resulting from the pairwise externalities. The key of our analysis is a probabilistic argument showing that these dependencies can always be resolved in each iteration. We believe that this approach might be applicable to similar problems and is therefore of independent interest.

Our approach is flexible enough to extend the algorithm to more general settings. One such generalization is to incorporate resource-restrictions for players (non-symmetric players). We show that our 2-approximation algorithm for MAX-CG-POS-EXT can be adapted to the case where the resources available to each player are restricted. We also obtain an improved  $\frac{3}{2}$ -approximation algorithm when every player is restricted to two resources. The proof of the  $\frac{3}{2}$ -approximation factor crucially exploits a characterization of the extreme point solutions of the LP relaxation.

A natural extension of our model are *r-generalized congestion games* where each player  $i$  specifies externalities  $v_{iT_e}$  for all player subsets  $T \subseteq N \setminus \{i\}$  of size at most  $r$ . The utility function of player  $i$  for resource  $e$  and player set  $S \subseteq N$  is defined as  $u_{i,e}(S) = \sum_{T \subseteq S \setminus \{i\}, |T| \leq r} v_{iT_e}$ . Using this terminology, generalized congestion games correspond to 1-generalized congestion games. We extend our rounding procedure to *r-generalized congestion games* with positive externalities and derive an  $(r + 1)$ -approximation algorithm.

Finally, we settle a question left open by Blumrosen and Dobzinski [3]. The authors showed that an optimal assignment for player-specific congestion games with non-negative linear utility functions  $u_{i,e}(x) = a_{i,e}x$  can be computed efficiently. We show that this problem becomes NP-hard for affine utility functions  $u_{i,e}(x) = a_{i,e}x + b_{i,e}$ .

**Related Work.** There are various papers that study congestion games with negative or positive externalities. For example, negative externalities are studied in routing [14], scheduling and load balancing [2]. Positive externalities are studied in the context of cost sharing [7], facility location [1] and negotiations [5,6].

Meyers and Schulz [10] studied the complexity of finding a minimum cost solution in congestion games (according to Rosenthal's classical congestion game model [13]). They study several variants of the problem and prove NP-hardness results, as well as inapproximability results for some cases and polynomial time computability results for some other cases.

Chakrabarty, Mehta, Nagarajan and Vazirani [4] were the first to study player-specific congestion games from a centralized optimization perspective.

The authors study the cost-minimization variant of the problem where each player has a non-negative and non-decreasing cost function associated with each resource.<sup>2</sup> They show that computing an assignment of minimum total cost is NP-hard. The authors also derive some positive results for certain special cases of the problem (see [4] for details).

Most related to our work is the paper by Blumrosen and Dobzinski [3]. They study the problem of welfare maximization in player-specific congestion games with non-negative utility functions. Among other results, they give NP-hardness and inapproximability results for positive and negative externalities. They also provide a randomized 18-approximation algorithm for arbitrary (non-negative) utility functions.

The problem of computing a welfare maximizing assignment for generalized congestion games can also be interpreted as the following graph coloring problem: We are given a complete undirected graph on  $n$  vertices and  $m$  colors  $[m]$ . Every edge  $(i, j)$  (including self-loops) has a weight  $w_{ije} = v_{ije} + v_{jie}$  for each color  $e \in [m]$ . The goal is to assign a color to every node such that the total weight of all *monochromatic* edges, i.e., edges whose endpoints have the same color, is maximized. The weight of a monochromatic edge  $(i, j)$  is defined as  $w_{ije}$ , where  $e$  is the color of the endpoints. The minimization variant of this problem with identical weights  $w_{ije} = w_{ij}$  for all  $e \in [m]$  and every edge  $(i, j)$  is also known as the generalized graph coloring problem [9], graph  $k$ -partitioning [8], and  $k$ -min cluster [15].

## 2 Preliminaries

A *generalized congestion game* is given by  $\Gamma = (N, E, \{\Sigma_i\}_{i \in N}, \{u_{i,e}\}_{i \in N, e \in E})$ , where  $N = [n]$  is the set of players,  $E = [m]$  is the set of facilities (or resources),  $\Sigma_i \subseteq 2^E$  is the strategy set of player  $i$ , and  $u_{i,e} : 2^N \rightarrow \mathbb{R}$  is the utility function that player  $i$  associates with facility  $e$ . Unless stated otherwise, we assume throughout this paper that  $\Sigma_i = E$  for every player  $i$  (symmetric players, singleton resource sets). The set  $\Sigma = \times_{i=1}^n \Sigma_i$  is called the *joint strategy set*, and elements in it are *strategy profiles*. Each player chooses a strategy  $s_i \in \Sigma_i$  from his strategy set, which gives rise to a strategy profile  $s = (s_1, \dots, s_n) \in \Sigma$ . A player tries to maximize his utility  $u_i(s) = u_{i,s_i}(\{j : s_j = s_i\})$ .

The utility functions  $u_{i,e}$  that we consider in this paper are of a specific form: Each player  $i$  associates with each facility  $e$  and player  $j$  a non-negative value  $v_{ije}$ . Then  $u_{i,e}$  is given by  $S \mapsto \sum_{j \in S} v_{ije}$ . The idea behind this is that for  $(i, j) \in N^2$  and  $e \in E$ ,  $v_{ije}$  specifies player  $i$ 's value of being on facility  $e$  together with player  $j$ . Here,  $v_{iie}$  is the value of player  $i$  for facility  $e$ , independently of the other players. We speak of a generalized congestion game with *positive*, *negative* or *mixed externalities*, respectively, when the  $v_{ije}$ 's are positive, negative or arbitrary. Note that allowing players to remain unassigned would not result in a more expressive model, as this is equivalent to assigning a player to an artificial facility for which the player always has utility 0.

<sup>2</sup> Equivalently, the utility functions are assumed to be non-positive and non-increasing.

Associated to  $\Gamma$  is a *social welfare function*  $\Pi : \Sigma \rightarrow \mathbb{R}$  defined by  $s \mapsto \sum_{i=1}^n u_i(s)$ . We are interested in the problem of finding an *optimal* strategy profile for a given generalized congestion games  $\Gamma$ , i.e., a strategy profile  $s^* \in \Sigma$  that maximizes  $\Pi$ . We use MAX-CG-XXX-EXT as a short to refer to this optimization problem, where  $\text{xxx} \in \{\text{POS}, \text{NEG}, \text{MIX}\}$  indicates the respective type of externalities.

Due to space limitations, many proofs are omitted and will be included in an extended version of this paper.

### 3 Mixed Externalities

We start off by studying the problem MAX-CG-MIX-EXT of optimizing the social welfare in generalized congestion games with mixed externalities. It turns out that this problem is highly inapproximable, even for 2 facilities.

**Theorem 1.** *MAX-CG-MIX-EXT is strongly NP-hard and is not  $n^{1-\epsilon}$ -approximable in polynomial time for every  $\epsilon > 0$ , unless  $\text{P} = \text{NP}$ , even for  $m = 2$  facilities.*

As it turns out, the problem can be solved efficiently if the number of players is fixed.

**Proposition 1.** *MAX-CG-MIX-EXT can be solved in polynomial time for a fixed number of players.*

### 4 Positive Externalities

Given the strong inapproximability result for mixed externalities (Theorem 1), we focus on the case of positive externalities in the remainder of this paper. Central to our study of this problem is the following integer program (IP) for the problem:

$$\max \sum_{e=1}^m \left( \sum_{i=2}^n \sum_{j=1}^{i-1} (v_{ije} + v_{jie}) x_{\{i,j\},e} + \sum_{i=1}^n v_{iie} x_{i,e} \right) \quad (1)$$

$$\text{s.t.} \quad \sum_{e=1}^m x_{i,e} = 1 \quad \forall i \in N \quad (2)$$

$$x_{\{i,j\},e} - x_{i,e} \leq 0 \quad \forall i, j \in N, \forall e \in E \quad (3)$$

$$x_{\{i,j\},e} \in \{0, 1\} \quad \forall i, j \in N, \forall e \in E \quad (4)$$

Recall that  $N = [n]$  and  $E = [m]$ . The variables are interpreted as follows:  $x_{i,e}$  is the 0-1 variable that indicates whether player  $i \in N$  is assigned to resource  $e$ ;  $x_{\{i,j\},e}$  is the 0-1 variable that indicates whether both players  $i, j \in N$ ,  $i \neq j$ , are assigned to resource  $e \in E$ . There are thus  $(n^2 + n)/2$  variables in total. Note that we exploit in the above formulation that all externalities are positive.

In the LP relaxation the constraints (4) are replaced by “ $0 \leq x_{\{i,j\},e} \leq 1, \forall i, j \in N, \forall e \in E$ ”. We can show that the LP relaxation of IP (1) is totally unimodular for  $m = 2$ . This implies the following positive result, which is in stark contrast with Theorem 1.

**Theorem 2.** MAX-CG-POS-EXT can be solved in polynomial time for  $m = 2$  facilities.

Unfortunately, the problem becomes strongly NP-hard for  $m \geq 3$ .

**Theorem 3.** MAX-CG-POS-EXT is strongly NP-hard for  $m \geq 3$  facilities.

#### 4.1 A 2-Approximate LP-Rounding Algorithm

In this section, we derive a 2-approximation algorithm for MAX-CG-POS-EXT. We first need to introduce some more notation.

We extend  $\Pi$  to the domain of fractional solutions of the LP relaxation of IP (1) and use the term *social welfare* to refer to the objective function of IP (1).

For a congestion game  $\Gamma = (N, E, \{\Sigma_i\}_{i \in N}, \{u_{i,e}\}_{i \in N, e \in E})$  with positive externalities  $\{v_{ije}\}_{i,j \in N, e \in E}$ , we define for  $e \in E$  and  $a \in \mathbb{Q}_{\geq 1}$  the  $(e, a)$ -boosted game as the game obtained by introducing a copy  $e'$  of facility  $e$  which all players value  $a$  times as much as the original  $e$ . Formally, the  $(e, a)$ -boosted game of  $\Gamma$  is the game  $\Gamma' = (N, E \cup \{e'\}, \{\Sigma_i \cup \{e'\}\}_{i \in N}, \{u_{i,f}\}_{i \in N, f \in E \cup \{e'\}})$  where  $u_{i,e'}(S) = \sum_{j \in S} av_{ije}$  for all  $i \in N$ . In an  $(e, a)$ -boosted game, we refer to the introduced facility  $e'$  as *the boosted facility*.

We fix for each facility  $e \in E$  a total order  $\preceq_e$  on  $N$  which satisfies  $i \prec_e j$  whenever  $x_{ie} < x_{je}$ . Using  $\preceq_e$ , we define  $P(e, k)$  as the player set  $N'$  for which it holds that  $|N'| = k$  and  $i \in N'$  if  $j \in N'$  and  $i \succ_e j$ . Informally,  $P(e, k)$  consists of the  $k$  players with the highest fractional assignments on  $e$ .

Finally, for a fractional solution  $x$  for  $\Gamma$ , we define the  $(e, a, k)$ -boosted assignment as the fractional solution to the  $(e, a)$ -boosted game where players  $P(e, k)$  are assigned integrally to the boosted facility and the remaining players are assigned according to  $x$ .

Our algorithm, BOOST( $a$ ), is a rounding algorithm that takes as its starting point the fractional optimum of the relaxation of IP (1), and iteratively picks some facility  $e \in E$  and assigns a set of players to the boosted facility in the  $(e, a)$ -boosted game. The formal description of BOOST( $a$ ) is given in Algorithm 1.

**Theorem 4.** Algorithm BOOST(2) is a deterministic polynomial time 2-approximation algorithm for MAX-CG-POS-EXT.

*Proof.* The hard part of the proof is to show that in each iteration there exist  $e$  and  $k$  such that Step 2.1 is feasible. This is done in Lemma 1 given below.

The algorithm clearly outputs a feasible solution. It is straightforward to check that the algorithm runs in polynomial time: For Step 1, solving an LP to optimality can be done in polynomial time using the ellipsoid method or an interior point method. For Step 2, there are only  $nm$   $(e, k)$ -pairs to check. Lastly,

in each iteration, at least one player will be assigned to a boosted facility, hence removed from  $[m]$ , so there are at most  $n$  iterations.

It is also easy to prove that Algorithm 1 outputs a solution for which the social welfare is within a factor  $\frac{1}{2}$  from the optimal social welfare: In the solution  $x'$  at the beginning of Step 3, every player  $i \in [n]$  is assigned to a copy  $e'$  of a facility  $e \in [m]$  for which it holds that  $v_{ije'} = 2v_{ije}$  for all  $j$ . So by assigning all players on such facilities  $e'$  to the original facility  $e$  decreases the social welfare at most factor 2. If we denote by  $SOL$  the social welfare of the strategy profile found by Algorithm 1, and if we denote by  $OPT$  the optimial social welfare, then the following sequence of inequalities proves that the algorithm outputs a 2-approximate solution:<sup>3</sup>

$$SOL = \Pi_{\Gamma}(x'') \geq \frac{1}{2} \Pi_{\Gamma'}(x') \geq \frac{1}{2} \Pi_{\Gamma}(x) \geq \frac{1}{2} OPT.$$

□

We come to the proof of the following lemma.

**Lemma 1.** *Suppose that  $x'$  is a solution to IP (1) for a congestion game  $\Gamma'$  with positive externalities. Denote by  $E$  the facility set of  $\Gamma'$  and assume that  $E = [m] \cup E'$ , where  $E'$  is a set of facilities such that  $x_{\{i,j\},e'} \in \{0,1\}$  for all  $i, j \in [n], e' \in E'$ . Then there is a facility  $e \in [m]$  and a number  $k \in [n]$  such that the social welfare of the  $(e, 2, k)$ -boosted assignment is at least the social welfare of  $x'$ . Moreover,  $e$  and  $k$  can be found in polynomial time.*

For the proof of Lemma 1, we need the following technical result:

**Lemma 2.** *Let  $a_1, a_2, \dots, a_n \in \mathbb{R}_{\geq 0}$ , be a non-increasing sequence of non-negative numbers with  $a_1$  positive, and let  $b_1, \dots, b_n \in \mathbb{R}$ . Suppose that  $\sum_{i=1}^n a_i b_i \geq 0$ . Then, there is a  $k \in [n]$  such that  $\sum_{i=1}^k b_i \geq 0$ .*

*Proof.* Let  $n'$  be the highest index such that  $a_{n'} > 0$ . There are two cases: either there is a  $k < n'$  for which the claim holds, or there is not such a  $k$ . For the latter case, we show that the claim must hold for  $k = n'$ . It follows from the following derivation:

$$0 \leq \sum_{i=1}^{n'} a_i b_i = \sum_{j=1}^{n'-1} (a_j - a_{j+1}) \sum_{i=1}^j b_i + a_{n'} \sum_{i=1}^{n'} b_i \leq a_{n'} \sum_{i=1}^{n'} b_i.$$

□

*Proof (of Lemma 1).* It suffices to only show existence of the appropriate  $e$  and  $k$ , as finding them in polynomial time can then simply be done by complete enumeration of all  $(e, k)$ -pairs (because there are only  $mn$  such pairs).

<sup>3</sup> The function  $\Pi$  is subscripted with the game of which it is the social welfare function.

---

**Algorithm 1** BOOST( $a$ ): An LP rounding algorithm for MAX-CG-POS-EXT.

---

1. Solve the relaxation of IP (1) for congestion game  $\Gamma$ , and let  $x$  be the optimal fractional solution. Let  $\Gamma' := \Gamma$  and  $x' = x$ .
  2. Repeat the following until  $x'$  is integral:
    - 2.1. Find a facility  $e \in [m]$  and a number  $k \in [n]$  such that the social welfare of the  $(e, a, k)$ -boosted assignment is at least the social welfare of  $x'$ .
    - 2.2. Let  $\Gamma'$  be the  $(e, a)$ -boosted game and let  $x'$  be the  $(e, a, k)$ -boosted assignment.
  3. For every player  $i \in [n]$  let  $e_i$  be the facility of  $[m]$  such that  $i$  is assigned to a boosted copy of  $e_i$  in  $x'$ . Let  $x''$  be the integral solution of IP (1) for  $\Gamma'$  obtained by assigning every  $i$  integrally to the original non-boosted facility  $e_i$ . Output the strategy profile for  $\Gamma$  that corresponds to  $x''$ .
- 

For  $e \in E$  and  $k \in [n]$ , let  $\Delta(e, k)$  denote the amount by which social welfare increases when comparing the  $(e, 2, k)$ -boosted assignment of  $x'$  to the  $(e, 2, k - 1)$ -boosted assignment of  $x'$ . Let  $p(e, k)$  be the single player in  $P(e, k) \setminus P(e, k - 1)$ . We can express  $\Delta(e, k)$  as  $\Delta^+(e, k) - \Delta^-(e, k)$ , where  $\Delta^+(e, k)$  is the increase in social welfare due to the additional utility on the boosted facility, and  $\Delta^-(e, k)$  is the loss in utility due to setting the assignment for player  $p(e, k)$  to 0 on all facilities in  $[m]$ . For notational convenience, we define  $w_{\{i,j\},e} = v_{ije} + v_{jie}$  and  $w_{i,e} = v_{iie}$ . Then we can write  $\Delta^+(e, k)$  and  $\Delta^-(e, k)$  as follows:

$$\begin{aligned} \Delta^+(e, k) &= 2w_{p(e,k),e} + \sum_{j:j \succ_e p(e,k)} 2w_{\{p(e,k),j\},e} \\ \Delta^-(e, k) &= \sum_{f \in [m]} \left( x_{p(e,k),f} w_{p(e,k),f} + \sum_{j:j \prec_e p(e,k)} x_{\{p(e,k),j\},f} w_{\{p(e,k),j\},f} \right). \end{aligned}$$

Clearly, if we move for some  $e \in [m]$  and  $k \in [n]$  the players  $P(e, k)$  to the boosted facility  $e'$  in the  $(e, 2)$ -boosted game, then the change in utility is  $\sum_{i=1}^k \Delta(e, i)$ . We therefore need to show that there is a facility  $e \in E$  and  $k \in [n]$  such that  $\sum_{i=1}^k \Delta(e, i) \geq 0$ .

To show this, let  $X$  be a random variable that takes on the values  $\{\Delta(e, k) : e \in [m], k \in [n]\}$ , of which the distribution is given by

$$\Pr[X = \Delta(e, k)] = \frac{x_{p(e,k),e}}{\sum_{e \in [m], i \in [n]} x_{p(e,i),e}} \quad \forall e \in [m], k \in [n].$$

Define  $Y = \sum_{e \in [m], i \in [n]} x_{p(e,i),e}$ .

We derive the following bound on the expectation of  $X$ :

$$\begin{aligned} \mathbf{E}[X] &= \sum_{(e,k): e \in [m], k \in [n]} \Pr[X = \Delta(e, k)] \Delta(e, k) \\ &= \frac{1}{Y} \sum_{e \in [m], k \in [n]} x_{p(e,k),e} \left( 2w_{p(e,k),e} + \sum_{j:j \succ_e p(e,k)} 2w_{\{p(e,k),j\},e} \right) \end{aligned}$$



$$\begin{aligned}
& - \sum_{f \in [m]} \left( x_{p(e,k),f} w_{p(e,k),f} + \sum_{j: j \prec_e p(e,k)} x_{\{p(e,k),j\},f} w_{\{p(e,k),j\},f} \right) \\
= & \frac{1}{Y} \left( \sum_{e \in [m], k \in [n]} 2x_{p(e,k),e} w_{p(e,k),e} \right. \\
& + \sum_{e \in [m], k \in [n]} \sum_{j: j \succ_e p(e,k)} 2x_{p(e,k),e} w_{\{p(e,k),j\},e} \\
& - \sum_{f \in [m], k \in [n]} \left( \sum_{e \in [m]} x_{p(e,k),e} \right) x_{p(e,k),f} w_{p(e,k),f} \\
& - \sum_{e \in [m], k \in [n]} x_{p(e,k),e} \sum_{f \in [m]} \left( \sum_{j: j \prec_e p(e,k)} x_{\{p(e,k),j\},f} w_{\{p(e,k),j\},f} \right) \Big) \\
= & \frac{1}{Y} \left( \sum_{e \in [m], i \in [n]} 2x_{i,e} w_{i,e} + \sum_{\{i,j\}, i \neq j} \sum_{e \in [m]} 2x_{\{i,j\},e} w_{\{i,j\},e} \right. \\
& - \sum_{e \in [m], i \in [n]} x_{i,e} w_{i,e} - \sum_{e \in [m], i \in [n]} x_{i,e} \sum_{f \in [m]} \sum_{j: j \prec_e i} x_{\{i,j\},f} w_{\{i,j\},f} \Big) \\
= & \frac{1}{Y} \left( \sum_{e \in [m], i \in [n]} x_{i,e} w_{i,e} + \sum_{\{i,j\}, i \neq j} \sum_{e \in [m]} 2x_{\{i,j\},e} w_{\{i,j\},e} \right. \\
& - \sum_{\{i,j\}, i \neq j} \left( \sum_{f \in [m]} x_{\{i,j\},e} w_{\{i,j\},e} \right) \left( \sum_{e \in [m]} \max\{x_{i,e}, x_{j,e}\} \right) \Big) \quad (5) \\
\geq & \frac{1}{Y} \sum_{e \in [m], i \in [n]} x_{i,e} w_{i,e} \geq 0.
\end{aligned}$$

In the above derivation we make use of the following facts that hold for all  $e \in [m], i, j \in [n]$ : (i)  $\sum_{f \in [m]} x_{i,f} = 1$  (for the fourth equality), (ii)  $x_{i,e} = \max\{x_{i,e}, x_{j,e}\}$  if  $i \succ_e j$  (for the fifth equality), (iii)  $\sum_{f \in [m]} \max\{x_{i,f}, x_{j,f}\} \leq 2$  (for the first inequality).

We can express  $\mathbf{E}[X]$  as the sum of the  $m$  terms  $\{T_e = \sum_{k \in [n]} \Delta(e, k) \Pr[X = \Delta(e, k)] : e \in [m]\}$ . Because the expectation is non-negative, it holds that  $T_e$  is non-negative for at least one  $e \in [m]$ . We take this  $e$  and apply Lemma 2 to  $T_e$  (take  $\Delta(e, i)$  for  $b_i$  and  $\Pr[X = \Delta(e, i)]$  for  $a_i$ ). We conclude that there is a  $k \in [n]$  such that  $\sum_{i=1}^k \Delta(e, k) \geq 0$ .  $\square$

## 4.2 A Matching Integrality Gap Example

In this section, we show that the integrality gap of IP (1) is very close to 2. This implies that no algorithm based on the LP relaxation of IP (1) can achieve an approximation factor better than 2 and thus the approximation factor of BOOST(2) is essentially best possible.

We construct a family of generalized congestion games, parameterized by the number of players  $n$ , for which we determine optimal integral and fractional solutions to IP (1) and its LP relaxation, respectively. It can be verified computationally that the integrality gap approaches 2 as  $n$  increases. We are convinced

that the integrality gap can be shown to approach 2 analytically, but at this point of time lack a formal proof of this.

Fix two parameters  $m$  and  $k$  with  $k \leq m$  and let  $H(m, k)$  be the  $k$ -uniform hypergraph on  $[m]$ . The instance  $I(m, k)$  is then defined as follows: there are  $m$  facilities and  $\binom{m}{k}$  players. We identify each player with a distinct edge of  $H(m, k)$ . The externality  $v_{ij e}$  is set to 1 if the hyperedge of  $i$  and the hyperedge of  $j$  both contain facility  $e$ , and 0 otherwise. The value  $v_{iie}$  is set to 1 if the hyperedge of  $i$  contains facility  $e$ .

We define  $OPT_{\text{frac}}(m, k)$  as the feasible fractional solution for  $I(m, k)$ , where each player is assigned with value  $1/k$  to each of its facilities. For each facility  $e \in [m]$  there are  $\frac{k}{m} \binom{m}{k} = \binom{m-1}{k-1}$  players that have  $e$  in their hyperedge and thus the social welfare of  $OPT_{\text{frac}}(m, k)$  is  $\frac{m}{k} \binom{m-1}{k-1}^2$ .

Using induction on  $m$ , one can straightforwardly prove that the social welfare of an optimal integral solution  $OPT_{\text{int}}(m, k)$  for this instance is  $\sum_{i=k-1}^{m-1} \binom{i}{k-1}^2$ . Subsequently, evaluating by computer the expression  $OPT_{\text{frac}}(m, k)/OPT_{\text{int}}$  for particular choices of  $m$  and  $k$  indicates that the integrality gap approaches  $\frac{2k-1}{k}$  as  $m$  gets larger. The largest integrality gap that we computed explicitly is 1.972013 for  $m = 5000$  and  $k = 71$ .

## 5 Variations on the Problem

We study in this section two generalizations and one special case of the problem.

**Restricted Strategy Sets.** Up to this point, we assumed that the strategy set  $\Sigma_i$  of each player  $i$  equals the facility set  $E$  (symmetric players, singleton strategy sets). We now show that our algorithm BOOST(2) generalizes to the case where every player  $i$  has a strategy set  $\Sigma_i \subseteq E$  (non-symmetric players, singleton strategy sets). The modification is easy: in Step 1, the algorithm solves a modified version of IP (1) with the additional constraints  $x_{i,e} = 0$  for all  $i \in N, e \notin \Sigma_i$ . The rounding procedure that follows after Step 1 of the algorithm then clearly puts every player  $i$  on a facility for which  $x_{i,e} \neq 0$ , so it produces a feasible integral solution. The rest of the proof of Theorem 4 remains valid.

**Corollary 1.** *There is a polynomial time 2-approximation algorithm for MAX-CG-POS-EXT with restricted strategy sets.*

For the special case that each player's strategy set is of size 2, we can improve the approximation factor.

**Proposition 2.** *There is a polynomial time  $\frac{3}{2}$ -approximation algorithm for MAX-CG-POS-EXT with restricted strategy sets of size 2.*

The proof of this proposition relies on a characterization result for the extreme point solutions of the LP relaxation of the modified IP, that we omit here. Moreover, one can show by a simple example that the integrality gap of the modified IP for strategy sets of size 2 is  $\frac{3}{2}$ .

**Externalities on Bigger Sets of Players.** Instead of restricting the players to express their externalities on pairs consisting of a single player and a facility, we can fix an arbitrary number  $r \in \mathbb{N}$  and allow players to express their externalities on pairs consisting of a set of at most  $r$  players and a facility. We refer to such congestion games as *r-generalized congestion games*. When each player  $i$  is allowed to express externalities  $v_{iT_e}$  for sets  $T \subseteq N \setminus \{i\}$ ,  $|T| \leq r$ , we will show that a simple adaptation of the algorithm  $\text{BOOST}(r+1)$  returns an  $(r+1)$ -approximate solution. The only change that needs to be made is that the relaxation of the following generalization of IP (1) is solved in Step 1 of  $\text{BOOST}(r+1)$ : For each player  $i$  and facility  $e$ , this generalized LP has again the variable  $x_{i,e}$  that indicates whether  $i$  is assigned to  $e$ . Moreover, for every facility  $e$  and set  $S$  of at most  $r+1$  players, there is a variable  $x_{S,e}$  that indicates whether all players in  $S$  are assigned to  $e$ . For notational convenience, let  $w_{S,e} = \sum_{i \in S} v_{i,S \setminus \{i\},e}$  for all sets  $S \subseteq N$ ,  $|S| \leq r$ . The generalized LP reads as follows:

$$\begin{aligned}
& \max \sum_{e=1}^m \sum_{S: S \subseteq [n], |S| \leq r+1} w_{S,e} x_{S,e} \\
& \text{s.t.} \quad \sum_{e=1}^m x_{i,e} \leq 1 && \forall i \in N \\
& \quad x_{S,e} - x_{i,e} \leq 0 && \forall S \subseteq [n], |S| \leq r+1, \forall i \in S, \forall e \in E, \\
& \quad x_{\{i,j\},e} \in \{0, 1\} && \forall i, j \in [n], \forall e \in E
\end{aligned}$$

We can show that the adapted version of  $\text{BOOST}(r+1)$  is a valid polynomial time  $(r+1)$ -approximation algorithm.

**Proposition 3.** *There is a polynomial time  $(r+1)$ -approximation algorithm for computing a social welfare maximizing strategy profile for  $r$ -generalized congestion games with positive externalities.*

**A Special Case: Affine Externalities.** In this section, we study the special case of our problem where the utility functions  $u_{i,e}(S)$  are *affine* non-negative functions of  $|S|$ , i.e., for every  $i \in N$  and each  $e \in E$ ,  $u_{i,e}(S) = a_{ie}|S| + b_{ie}$  for non-negative rational numbers  $a_{ie}, b_{ie}$ . It is straightforward to see that this type utility functions falls within our positive externalities model. We refer to the respective optimization problem as  $\text{MAX-CG-AFF-EXT}$ .

The motivation for studying this is that Blumrosen and Dobzinski show in [3] that if  $b_{ie} = 0$  for all  $i \in N$ ,  $e \in E$  then the optimal solution can be found in polynomial time. Allowing  $b_{ie}$  to be non-zero is thus one of the simplest generalizations that comes to mind.

In contrast to the polynomial time result in [3], we show that this problem is strongly  $\text{NP-hard}$ .

**Theorem 5.**  *$\text{MAX-CG-AFF-EXT}$  is strongly  $\text{NP-hard}$ .*

One could use the algorithm  $\text{BOOST}(2)$  to find a 2-approximate solution to this affine special case of the problem. However, it is easy to see that there is a

much simpler and faster 2-approximation algorithm for this case: The algorithm simply chooses the maximum among the strategy profiles  $s$  and  $t$  defined as follows: for  $i \in N$ ,  $s_i = \arg_e \max\{b_{ie} : e \in E\}$  and  $t_i = \arg_e \max\{\sum_{j \in N} na_{je} : e \in E\}$ .

It is shown in Proposition 4.3 of [3] that if  $b_{ie} = 0$  for all  $i \in N$ ,  $e \in E$ , then the strategy profile  $t$  is optimal. It is obvious that if  $a_{ie} = 0$  for all  $i \in N$ ,  $e \in E$ , then the strategy profile  $s$  is optimal. Therefore, for the case that neither  $a_{ie}$  nor  $b_{ie}$  is necessarily 0, the maximum of these two joint strategies must be within a factor of 2 from the optimal strategy profile.

## References

1. E. Anshelevich, A. Dasgupta, J. M. Kleinberg, E. Tardos, T. Wexler, and T. Roughgarden. The price of stability for network design with fair cost allocation. *SIAM Journal on Computing*, 38(4):1602–1623, 2008.
2. B. Awerbuch, Y. Azar, E.F. Grove, Ming-Yang Kao, P. Krishnan, and J.S. Vitter. Load balancing in the  $L_p$  norm. *IEEE Annual Symposium on Foundations of Computer Science*, 1995.
3. L. Blumrosen and S. Dobzinski. Welfare maximization in congestion games. In *Proceedings of the 7th ACM conference on Electronic commerce*, pages 52–61, 2006.
4. D. Chakrabarty, A. Mehta, V. Nagarajan, and V. Vazirani. Fairness and optimality in congestion games. In *Proceedings of the 6th ACM Conference on Electronic Commerce*, pages 52–57, New York, NY, USA, 2005. ACM.
5. V. Conitzer and T. Sandholm. Expressive negotiation over donations to charities. In *Proceedings of the 5th ACM conference on Electronic commerce*, EC '04, pages 51–60, New York, NY, USA, 2004. ACM.
6. V. Conitzer and T. Sandholm. Expressive negotiation in settings with externalities. In *Proceedings of the 20th national conference on Artificial intelligence - Volume 1*, AAAI'05, pages 255–260. AAAI Press, 2005.
7. J. Feigenbaum, C. H. Papadimitriou, and S. Shenker. Sharing the cost of multicast transmissions. *Journal of Computer and System Sciences*, 63(1):21–41, 2001.
8. V. Kann, S. Khanna, J. Lagergren, and A. Panconesi. On the hardness of approximating max  $k$ -cut and its dual. *Chicago Journal of Theoretical Computer Science*, 2, 1997.
9. A. W. J. Kolen and J. K. Lenstra. Combinatorics in operations research. In R. Graham, M. Grötschel, and L. Lovász, editors, *Handbook of Combinatorics*. Elsevier, Amsterdam, 1995.
10. C.A. Meyers and A.S. Schulz. The complexity of welfare maximization in congestion games. *Networks*, pages 252–260, 2012.
11. I. Milchtaich. Congestion games with player-specific payoff functions. *Games and Economic Behavior*, 13(1):111–124, 1996.
12. N. Nisan, T. Roughgarden, É. Tardos, and V. V. Vazirani, editors. *Algorithmic Game Theory*. Cambridge University Press, 2007.
13. R. W. Rosenthal. A class of games possessing pure-strategy nash equilibria. *International Journal of Game Theory*, 2(1):65–67, 1973.
14. T. Roughgarden. *Selfish Routing and the Price of Anarchy*. MIT Press, 2005.
15. S. Sahni and T. Gonzalez. P-complete approximation problems. *Journal of the ACM*, 23(3):555–565, 1976.