J Sched (2012) 15:157–164
DOI 10.1007/s10951-011-0247-8

Efficiency analysis of load balancing games with and without activation costs

Bo Chen · Sinan Gürel

Published online: 30 July 2011

Abstract In this paper, we study two models of resource allocation games: the classical load-balancing game and its new variant involving resource activation costs. The resources we consider are identical and the social costs of the games are utilitarian, which are the average of all individual players' costs.

Using the social costs we assess the quality of pure Nash equilibria in terms of the price of anarchy (PoA) and the price of stability (PoS). For each game problem, we identify suitable problem parameters and provide a parametric bound on the PoA and the PoS. In the case of the load-balancing game, the parametric bounds we provide are sharp and asymptotically tight.

Keywords Resource allocation game · Congestion cost · Load balancing · Cost sharing · Price of anarchy · Price of stability

1 Introduction

Problems of resource allocation often involve decentralized decision making. A typical example is, in terms of machine scheduling, allocation of machines to jobs (or assignment of jobs to machines) in which selfish agents, representing individual jobs, select machines for processing their own jobs. In the long run, decisions of the agents, motivated by individual interests, usually result in a *Nash equilibrium* (NE) at which no individual agent will benefit from any unilateral deviation for the current resource allocation. In terms of a given social objective, such an equilibrium is not necessarily, indeed can often be far from, optimal. It is important, therefore, to analyze the quality of NE solutions in terms of social optimality.

The resource allocation games we consider in this paper are as follows. Given a set of n jobs, each of which has a positive *length* and is controlled by a selfish agent. Each agent decides on which of the identical machines available to assign his job to. We consider two game models: the *load-balancing* model and the *cost-sharing* model.

In the load-balancing game, a fixed number m of machines are given and the cost of each player is caused by a congestion, which is defined as the load of the machine, i.e., the sum of lengths of the jobs assigned to it. This is the classical load-balancing game as surveyed in Vöcking (2007) and has been studied extensively. In our second game model there are unlimited number of identical machines available, but usage of each machine comes with an additional set-up or *activation* cost, which is shared in proportion to the job lengths by all agents who assign their jobs to the machine. This model has recently been introduced and studied in Feldman and Tamir (2008). As shown, respectively, in Vöcking (2007) and Feldman and Tamir (2008), any instance of the games in the two models admits a *pure* Nash equilibrium, which can be computed efficiently. We are concerned in this paper with the quality of pure Nash equilibria.

In assessing the quality of a resource allocation, most of the studies in the literature, including Vöcking (2007) and

B. Chen
Qufu Normal University, Shandong 273165, China

B. Chen (✉) · S. Gürel
Centre for Discrete Mathematics and its Applications (DIMAP), Warwick Business School, University of Warwick, Coventry, CV4 7AL, UK
e-mail: b.chen@warwick.ac.uk

S. Gürel
Department of Industrial Engineering, Middle East Technical University, 06800 Ankara, Turkey

Feldman and Tamir (2008) mentioned above, use as social cost the *maximum* (or ℓ_∞ metric) of all individual players' costs. In this study, we consider a *utilitarian* social objective, which is defined as the *average* (or ℓ_1 metric) of all individual players' costs. Such a social objective is also used in a number of other similar studies (such as Berenbrink et al. (2006), Hoefer and Souza (2011), Gairing et al. (2008b) and Roughgarden and Tardos (2002), in which infinitely many jobs are considered) and is a standard assumption in the multi-agent system literature (e.g., Endriss et al. 2003; McBurney et al. 2002; Sandholm 1998). As shown by Berenbrink et al. (2006), Nash equilibria behave very differently under ℓ_1 metric than ℓ_∞ metric.

As in many other games, the social cost of an NE solution in the resource allocation games is often not minimum, whose corresponding solutions are called *optimal*. In this paper, we use the commonly accepted notions of the *price of anarchy* (PoA) and the *price of stability* (PoS) to analyze the quality of NE solutions. As introduced in Koutsoupias and Papadimitriou (1999), Papadimitriou (2001), the PoA (respectively, PoS) is defined as the ratio of the social cost of the worst (respectively, best) NE solution and the corresponding optimal social cost. For the load-balancing game with social objective of ℓ_∞ metric, Koutsoupias and Papadimitriou (1999) prove initial bounds on the PoA, and Czumaj and Vöcking (2002) are the first to provide (asymptotically) tight bounds on the PoA for a general case where the resources are not necessarily identical but *related*, i.e., the machines have different *speeds* (known as uniform-machine environment). For more detailed coverage of related research on games of social costs with ℓ_∞ metric, the reader is referred to, e.g., articles of survey nature (Czumaj 2004; Vöcking 2007; Heydenreich et al. 2007) and the references therein.

In the load-balancing game, the utilitarian social cost has been considered in several studies in the literature. Berenbrink et al. (2006) consider pure strategies in the uniform-machine environment and show that, if all job lengths are at least 1, then PoA $\le 4p_{\max}$, where $p_{\max}$ is the maximum job length, and if *additionally* all machines are identical, then PoS $\ge \sqrt{p_{\max}}/5$. Therefore, it is variability of job lengths, as opposed to machine speeds, that may lead to a big PoS and PoA. We explore this issue in further depth in this paper by providing asymptotically tight (indeed, sharp) bounds on the PoS and the PoA in terms of average normalized job length under the identical-machine environment but without any restriction on job lengths.

Hoefer and Souza (2011) study the utilitarian social objective for a routing game of n players and m parallel links where each link has a different speed. They assume players can freely set individual message lengths. Gairing et al. (2008b) consider inter alia the utilitarian social objective in the routing game on identical links with incomplete information.

Resource allocation games or more generally weighted singleton congestion games of social costs similar to ours have been studied extensively in the literature. Most notable is the social cost of *total latency*, which is defined as the *weighed* total of all individual players' costs, where the weights are particularly the job lengths in our resource allocation games. Research on equilibrium quality in such games can be found, for example, in Awerbuch et al. (2005), Aland et al. (2006), Gairing and Scoppmann (2007), Gairing et al. (2008a), Lücking et al. (2008), Roughgarden (2009).

All the studies we have mentioned above are on games in which the cost functions of players are of the nature of negative congestion effect, that is, an individual cost incurred by using a resource is a non-decreasing function of its load. However, positive congestion effect also happens in situations where a resource user wishes to share the resource with as many additional users as possible to minimize his individual cost of using the resource, which is a non-increasing function of its load. Such games are considered in Feigenbaum et al. (2001), Anshelevich et al. (2008), Epstein et al. (2007) for studies of fair cost sharing in network routing and design. The second game model considered in this paper, which takes both congestion effects into account has been recently proposed by Feldman and Tamir (2008). With an egalitarian social objective (ℓ_∞ metric), they show that the PoA is unbounded but the PoS is bounded by a tight bound of $5/4$. In this paper, we use the utilitarian social objective (ℓ_1 metric) to study the performance of NE solutions in the game. We provide a tight parametric bound on the PoA, which is unbounded in general (as in the case of egalitarian social objective considered in Feldman and Tamir (2008)). Furthermore, we give a parametric *lower* bound on the PoS, which is unbounded in general, in contrast with the boundedness of the egalitarian social objective.

2 Model descriptions and main results

A set of jobs $J = \{1, 2, \dots, n\}$ is to be assigned to a number of identical parallel machines. Each job $j \in J$ has a length of $p_j > 0$. For a given job assignment $\mathcal{A}$, we denote the set (respectively, number) of jobs assigned to machine i by $J^{\mathcal{A}}[i]$ (respectively, $n_i^{\mathcal{A}}$). The load of machine i under assignment $\mathcal{A}$ is then $L_i^{\mathcal{A}} = \sum_{j \in J^{\mathcal{A}}[i]} p_j$. If the assignment $\mathcal{A}$ is optimal (w.r.t. the given utilitarian social objective), we use $J^*[i]$, n_i^* and L_i^* to denote the above quantities, respectively.

In the load-balancing model, we assume that there are a fixed number m of identical machines. In this model, the *cost* to a job is the load of the machine the job is assigned to. The social cost of a given overall job assignment $\mathcal{A}$ is

$$C_1(\mathcal{A}) = \sum_{i=1}^{m} n_i^{\mathcal{A}} L_i^{\mathcal{A}},$$

which is the sum of all individual job costs. Note that as far as the PoA and PoS are concerned, this (and the following) definition of the social cost is equivalent to the one with "sum" or "total" replaced by "average" (over all jobs).

In the cost-sharing model, let B denote the cost of activating each machine. Given an overall job assignment $\mathcal{A}$, if job j of length p_j is assigned to machine i of load $L_i^{\mathcal{A}}$, then the cost of the job is

$$L_i^{\mathcal{A}} + \frac{p_j}{L_i^{\mathcal{A}}} B,$$

where the first term is its resource usage cost and the second represents its share of the cost of activating machine i, which is in proportion to its length with respect to the total load of the machine. In this model, we define the social cost by

$$C_2(\mathcal{A}) = \sum_{i=1}^{m^{\mathcal{A}}} n_i^{\mathcal{A}} L_i^{\mathcal{A}} + m^{\mathcal{A}} B,$$

where $m^{\mathcal{A}}$ is the number of activated machines in the given job assignment $\mathcal{A}$. As can be seen easily, the cost $C_2(\mathcal{A})$ is the total of individual job costs.

In each π of the two game problems defined above ($\pi = 1, 2$), we will use C_π^e and C_π^* to denote the social cost of an NE and an optimal solution, respectively. Let $P = \sum_{j \in J} p_j$ and denote the minimum, average and maximum job lengths, respectively, by $p_{\min} = \min_{j \in J} p_j$, $p_{\text{avg}} = P/n$ and $p_{\max} = \max_{j \in J} p_j$. Let

$$\begin{cases} \rho_1 = p_{\text{avg}}/p_{\min}, \\ \rho_2 = B/p_{\min}. \end{cases} \tag{1}$$

We shall use ρ_1 and ρ_2 as parameters in bounding the PoA and PoS in the respective two game models we study. In load-balancing game, parameter ρ_1 represents the average job length normalized by $p_{\min}$. Note that, as we pointed earlier, in providing the upper bound of $4p_{\max}$ on the PoA, Berenbrink et al. (2006) assume that all job lengths are at least 1. This assumption has implicitly hidden the possibility of the unboundedness of the PoA when $p_{\min} \to 0$, which is indeed the case as we shall show. We take this issue into account by normalizing all job lengths with $p_{\min}$ in assessing the PoA and PoS. On the other hand, the sharpness of our parametric bounds in terms of ρ_1 demonstrates that it is the average (instead of maximum) job length that is more accurate and hence suitable in bounding the PoA and PoS. In the cost-sharing game, as we shall see, parameter ρ_2 adequately represents a maximum number of jobs that can be assigned to a machine in any NE solution. In the next two sections, we will establish the following two sets of main results for the load-balancing and cost-sharing models with utilitarian social objectives C_1 and C_2, respectively:

$$\rho_1 - 1 \le \text{PoS} \le \text{PoA} \le \rho_1 + 1$$

(Theorem 3.1: load-balancing model)

$$\frac{1}{4}(\sqrt{\rho_2} + 2) \le \text{PoS} \le \text{PoA} \le \frac{1}{2}(\rho_2 + 1)$$

(Theorem 4.1 & Corollary 4.2: cost-sharing model)

For the load-balancing model, the two bounds together show that they are very sharp and both asymptotically tight, in other words, the PoS and PoA are nearly the same. For the cost-sharing model, the bounds show that the PoS and PoA can be unbounded if the parameter ρ_2 is not restricted. The upper bound is derived from a *tight* bound but of *two* parameters presented in Theorem 4.1. Further discussions on the tightness of the lower and upper bounds are provided in the final section.

3 Load-balancing model

We start with a direct observation of a simple property of NE assignments. For notational simplicity we omit from now on the indication of any assignment $\mathcal{A}$ from the notation unless there is a confusion.

Observation 3.1 *Given any NE assignment, machine loads satisfy the following inequalities*:

$$L_i \le L_k + p_j, \quad \forall j \in J[i], 1 \le i, k \le m.$$

Observation 3.1 simply means that, in any NE assignment, no job can reduce its cost by unilaterally changing its machine. Using Observation 3.1 we next prove an upper bound on C_1^e.

Lemma 3.2 *Given any NE assignment, its total cost C_1^e satisfies the following*:

$$C_1^e = \sum_{i=1}^{m} n_i L_i \le \left(\frac{n}{m} + 1\right) P.$$

Proof For any fixed i ($1 \le i \le m$), we choose k and j in Observation 3.1 such that

$$L_k = \min_{1 \le k' \le m} L_{k'} \le \frac{P}{m}; \quad \text{and} \quad p_j = \min_{j' \in J[i]} p_{j'} \le \frac{L_i}{n_i}.$$

The two inequalities above are evident. Therefore, we obtain

$$L_i \le \frac{P}{m} + \frac{L_i}{n_i}, \quad i = 1, \ldots, m,$$

which leads directly to our conclusion. □

The upper bound on C_1^e in Lemma 3.2 depends on the total length of the jobs, the number of jobs and the number of machines. The following lemma is a direct conclusion from the convexity of function $f(x) = x^2$.

Lemma 3.3 *For any real values* $x_1, \dots, x_m$, *we have*

$$\sum_{i=1}^{m} x_i^2 \geq \frac{1}{m}\left(\sum_{i=1}^{m} x_i\right)^2.$$

With Lemmas 3.2 and 3.3, we can establish our first upper bound on the PoA.

Theorem 3.1 *Let* ρ_1 *be defined as in* (1). *Then*

$$\rho_1 - 1 \leq PoS \leq PoA \leq \rho_1 + 1.$$

Proof Fix any NE and optimal assignments. From Lemma 3.2 and the fact that $P = \sum_i L_i^* \leq C_1^*$, we have

$$\frac{C_1^e}{C_1^*} \leq 1 + \frac{n}{m} \frac{P}{\sum_{i=1}^{m} n_i^* L_i^*}.$$

With the following inequality

$$L_i^* \geq n_i^* \, p_{\min},$$

and Lemma 3.3 and noticing that $\sum_{i=1}^{m} n_i^* = n$, we get

$$\frac{C_1^e}{C_1^*} \leq 1 + \frac{n}{m} \frac{P}{p_{\min} \sum_{i=1}^{m} (n_i^*)^2} \leq \rho_1 + 1.$$

The following example provides the lower bound for the PoS. □

Example 3.4 Consider an instance of m machines, $m-1$ large jobs of unit length and n small jobs of length $1/n$. Assume that $n > m$ and let n be a multiple of $m(m-1)$. Consider the assignment in which all large jobs are assigned to a single machine and all small jobs are evenly assigned to the remaining machines. The social cost of this assignment is an upper bound on the optimum:

$$C_1^* \leq (m-1)^2 + \frac{n}{m-1}.$$

Now consider the NE assignment $\mathcal{A}$ in which one large job is assigned to each of $(m-1)$ machines and all small jobs are assigned to the last machine. The social cost of the assignment is

$$C_1^e(\mathcal{A}) = (n + m - 1).$$

Then we get

$$\frac{C_1^e(\mathcal{A})}{C_1^*} \geq \frac{(n+m-1)(m-1)}{(m-1)^3 + n},$$

which approaches $m-1$ as $n \to +\infty$. There are no NE assignments other than $\mathcal{A}$. Note that

$$\rho_1 = \frac{m/(n+m-1)}{1/n},$$

which approaches m as $n \to \infty$. Therefore, the PoS is bounded from below by $\rho_1 - 1$.

Remark A special case of Proposition 4.2 in Gairing et al. (2008b) would lead to a slightly improved upper bound in Lemma 3.2 with n replaced by $n-1$, which would in turn contribute to a reduction of the upper bound in Theorem 3.1 to $\rho_1 + (n-1)/n$. However, this does not help as the number n of jobs can be arbitrarily large. We have decided to use Lemma 3.2 as its proof is much simpler.

4 Cost-sharing model

Recall that in the cost-sharing model, there are unlimited number of machines available, but usage of each machine incurs an additional activation cost of B. As we shall see, it is convenient to divide the jobs into two categories, large and small: $J_l = \{j \in J : p_j > B\}$ and $J_s = \{j \in J : p_j \leq B\}$, and the problem becomes trivial if all jobs are large: $J_s = \emptyset$. For notational convenience and without loss of generality, we will assume $B = 1$ in the remainder of this section, as we can achieve this by dividing all job lengths with the activation cost B.

It is easy to observe the following property of NE assignments for large jobs.

Lemma 4.1 *Any large job will be assigned to a dedicated machine in any NE assignment.*

Proof Consider a large job with a length $p > 1$. Suppose that there exists an NE in which the large job shares a machine with some other jobs that have a total length of q. Then the individual cost of the large job is

$$p + q + \frac{p}{p+q}$$

which is greater than $p+1$, its individual cost on a dedicated machine. This is because

$$p + q + \frac{p}{p+q} - p - 1 = q - \frac{q}{p+q} > 0$$

since $1/(p+q) < 1$. Then in an NE assignment a large job has to be on a dedicated machine. □

Similarly, the following lemma characterizes optimal assignments.

Lemma 4.2 *In any optimal assignment, if $L_i^* > 1$, then $n_i^* = 1$.*

Proof Suppose that $L_i^* > 1$ and $n_i^* \geq 2$. Let $j \in J^*[i]$. Then moving job j from machine i to a dedicated machine will result in a new assignment with a reduced objective value C_2':

$$\begin{aligned} C_2' - C_2^* &= \big(1 + (n_i^* - 1)(L_i^* - p_j) + 1 + p_j\big) \\ &\quad - \big(1 + n_i^* L_i^*\big) \\ &= 1 - L_i^* - (n_i^* - 2)p_j < 0, \end{aligned}$$

which contradicts the optimality of the original schedule. □

Lemma 4.2 implies that an optimal assignment assigns all large jobs to dedicated machines. On NE assignments of small jobs only, we have

Lemma 4.3 *For machines of small jobs, $L_i \leq 1$ holds in any NE assignment.*

Proof Suppose that $L_i > 1$ holds for machine i. Consider job j on machine i. The cost of job j is

$$L_i + \frac{p_j}{L_i}.$$

If job j activates a new machine, its cost will be $1 + p_j$. Then the cost change is

$$\begin{aligned} \Delta &= 1 + p_j - L_i - \frac{p_j}{L_i} = (1 - L_i) + p_j \frac{L_i - 1}{L_i} \\ &= \frac{1 - L_i}{L_i}(L_i - p_j) < 0. \end{aligned}$$ □

Lemma 4.3 implies that in an NE assignment, no machines other than dedicated ones can have a load greater than 1. Given Lemmas 4.1 and 4.3, let $A = \sum_{j \in J_l} p_j + |J_l|$ and denote $n_s = |J_s|$ and $P_s = \sum_{j \in J_s} p_j$, we have the following result:

Lemma 4.4 *Any NE assignment has a social cost C_2^e that is bounded from above as follows:*

$$C_2^e \leq n_s + P_s + A = n + P.$$

The bound is tight if $L_i = 1$ for all i.

Proof Suppose that there are m activated machines in the NE assignment. According to Lemma 4.1, we assume that all small jobs are assigned to the first m_s machines. Then $L_i \leq 1$ for $i = 1, \ldots, m_s$. We have

$$\begin{aligned} C_2^e &= \sum_{i=1}^{m_s} (n_i L_i + 1) + A \\ &= \sum_{i=1}^{m_s} \big((n_i - 1)L_i + 1\big) + P_s + A \\ &\leq n_s + P_s + A. \end{aligned}$$

The statement on tightness is easy to verify. □

On the other hand, let us provide a lower bound on C_2^*. Note that Lemmas 4.1 and 4.2 together imply that, if $J_s = \emptyset$, then any NE assignment is optimal and vice versa. Therefore, without loss of generality, we assume $J_s \neq \emptyset$ and define

$$\tau = \frac{B}{\max\{p_j : j \in J_s\}} = \frac{1}{\max\{p_j : j \in J_s\}}. \tag{2}$$

Lemma 4.5 *Any optimal assignment has a social cost C_2^* such that*

$$C_2^* \geq 2P_s\sqrt{\tau} + A.$$

Proof Since $n_i^* \geq \tau L_i^*$, we have

$$C_2^* = \sum_{i=1}^{m_s^*} n_i^* L_i^* + m_s^* + A \geq \sum_{i=1}^{m_s^*} \tau L_i^{*2} + m_s^* + A.$$

Since $\sum_{i=1}^{m_s^*} L_i^* = P_s$, from Lemma 3.3 we have $\sum_{i=1}^{m_s^*} L_i^{*2} \geq P_s^2/m_s^*$. Then,

$$C_2^* \geq \tau \frac{P_s^2}{m_s^*} + m_s^* + A.$$

The fact that the right-hand side of the above inequality is at its minimum when $m_s^* = P_s\sqrt{\tau}$ implies

$$C_2^* \geq 2P_s\sqrt{\tau} + A.$$ □

With the above lemma, we are now able to establish our second main result.

Theorem 4.1 *Let ρ_2 be defined as in (1) and τ be defined as in (2). Then the PoS and PoA are bounded as follows:*

$$\frac{1}{4}(\sqrt{\rho_2} + 2) \leq PoS \leq PoA \leq \frac{\rho_2 + 1}{2\sqrt{\tau}}.$$

Furthermore, the upper bound is tight.

Since $\tau \geq 1$, we immediately have the following:

Corollary 4.2 *The upper bound on the PoA in Theorem 4.1 can be simplified to contain only one parameter:*

$$PoA \leq \frac{\rho_2 + 1}{2}.$$

Proof of Theorem 4.1 Given the upper bound on C_2^e and the lower bound on C_2^* in Lemmas 4.4 and 4.5, respectively, noticing that $n_s \le \rho_2 P_s$, we obtain

$$\frac{C_2^e}{C_2^*} \le \frac{n_s + P_s + A}{2P_s\sqrt{\tau} + A} \le \frac{\rho_2 + 1}{2\sqrt{\tau}}.$$

Since the above inequality holds for any instance, the upper bound in the theorem is established. The tightness of the upper bound and validity of the lower bound are shown in the following two examples, respectively. □

Example 4.6 Consider an instance of a^{2k} jobs, each having a length of $1/a^{2k}$, where $a, k > 1$ are fixed integers. An NE assignment is that all jobs are on a single machine, which has the following social cost:

$$C_2^e = a^{2k} + 1.$$

It is easy to see that an optimal solution distributes all these identical jobs evenly on all activated machines, i.e., $|n_u^* - n_v^*| \le 1$ for any two activated machines u, v. Suppose that m machines are activated in a solution and there are equal numbers of jobs on all these machines. Then the social cost of such a solution is

$$C_2 = m + \frac{a^{2k}}{m}\frac{1}{a^{2k}}a^{2k} = m + \frac{a^{2k}}{m},$$

which is minimized if $m^* = a^k$. Hence, there exists an optimal (also NE) solution with a^k machines activated and of the following social cost:

$$C_2^* = 2a^k.$$

Therefore, we conclude that

$$\frac{C_2^e}{C_2^*} = \frac{a^{2k}+1}{2a^k} = \frac{\rho_2 + 1}{2\sqrt{\tau}},$$

which equals the upper bound in Theorem 4.1.

Example 4.7 Consider an instance of $n+1$ jobs ($n \ge 2$) with machine activation cost $B = 1$. The job lengths are $p_j = 1/n^2$ for $j = 1, \dots, n$ and $p_{n+1} = 1 - 1/n - \epsilon$ with $0 < \epsilon < 1/n^2$. Let $\mathcal{S}$ be the assignment in which all jobs are assigned to the same machine, say machine 1. Then $L_1 = 1 - \epsilon$, giving a social cost

$$\begin{aligned} C_2 = C_2(\mathcal{S}) &= n_1 L_1 + 1 = (n+1)(1-\epsilon) + 1 \\ &= n + 2 - (n+1)\epsilon. \end{aligned}$$

Proposition 1 *$\mathcal{S}$ is the unique NE assignment.*

Proof Assignment $\mathcal{S}$ is an NE assignment since $L_1 = 1 - \epsilon$, no job can reduce its cost by activating a new machine. We show that $\mathcal{S}$ is the unique NE assignment. We shall often use the fact that real function

$$f_r(x) = x + \frac{r^2}{x}, \quad x > 0\ (r > 0 \text{ fixed})$$

is convex and strictly decreasing over $(0, r]$ and strictly increasing over $[r, +\infty)$.

Claim 1 *In any NE assignment with at least two machines activated, job $n+1$ cannot have a dedicated machine.*

Suppose that job $n+1$ is on a dedicated machine, say, machine 1. Then the cost of job $n+1$ is $p_{n+1} + 1$. On the other hand, there is another machine i of load L_i ($i \ne 1$) such that

$$\frac{1}{n^2} \le L_i \le \frac{1}{n}. \tag{3}$$

Then job $n+1$ would benefit by deviating from machine 1 to machine i since

$$L_i + p_{n+1} + \frac{p_{n+1}}{L_i + p_{n+1}} < p_{n+1} + 1, \tag{4}$$

which can be easily verified by noticing that $p_{n+1} < 1 - 1/n$ and the fact that the left-hand side of inequality (4) is convex in L_i and hence maximized at the endpoints (in fact, the right endpoint) of interval (3).

Claim 2 *In any NE assignment with at least two machines activated, no job will share the same machine with job $n+1$.*

Suppose to the contrary that such NE assignment exists and a (nonempty) subset of jobs from $1, \dots, n$ are on machine 1, which also contains job $n+1$. Then, the individual cost of any job on machine 1 other than job $n+1$ is

$$\Gamma_1(L_1) \equiv L_1 + \frac{1/n^2}{L_1}, \quad \text{where}$$

$$L_1 \ge 1 - \frac{1}{n} - \epsilon + \frac{1}{n^2} > 1 - \frac{1}{n} \equiv \lambda_1.$$

This cost is strictly increasing in L_1 and thus is *more than* $\Gamma_1(\lambda_1)$.

Consider moving a job on machine 1 other than job $n+1$ to another activated machine i. Then the individual cost of the moving job will become

$$\Gamma_2(L_i) \equiv L_i + \frac{1}{n^2} + \frac{1/n^2}{L_i + 1/n^2}, \quad \text{where}$$

$$\lambda_2 \equiv \frac{1}{n^2} \le L_i \le \frac{1}{n} - \frac{1}{n^2},$$

which is strictly decreasing in L_i and thus achieves its maximum at $L_i = \lambda_2$. Since

$$\Gamma_1(\lambda_1) = 1 - \frac{n-2}{n(n-1)} \quad \text{and}$$

$$\Gamma_2(\lambda_2) = \frac{1}{2} + \frac{2}{n^2}$$

it is easily checked that $\Gamma_1(\lambda_1) \geq \Gamma_2(\lambda_2)$, which implies that the moving job will benefit from such a unilateral deviation, contradicting that the original assignment is an NE.

Claims 1 and 2 together imply that no NE assignment will activate more than one machine, which proves our proposition. □

Let $\mathcal{S}^*$ be the assignment in which machine 1 is dedicated to job $n+1$ and machine 2 accommodates all other jobs $1, \dots, n$. The social cost of $\mathcal{S}^*$ is

$$C_2(\mathcal{S}^*) = \sum_{i=1,2} n_i L_i + 2 = 4 - \frac{1}{n} - \epsilon.$$

Proposition 2 *$\mathcal{S}^*$ is an optimal assignment.*

Proof An assignment with at least three machines activated will have a social cost more than $3 + 1 - 1/n - \epsilon$. □

Consequently, with Propositions 1 and 2, the lower bound on the PoS in Theorem 4.1 is implied by the fact that, as $\epsilon \to 0$, we have

$$\frac{C_2(\mathcal{S})}{C_2(\mathcal{S}^*)} = \frac{n+2-(n+1)\epsilon}{4 - \frac{1}{n} - \epsilon} \to \frac{n+2}{4 - \frac{1}{n}}$$

$$> \frac{n+2}{4} = \frac{1}{4}\left(\sqrt{\rho_2} + 2\right).$$

Therefore, for ϵ sufficiently small, the bound is valid.

5 Concluding remarks and further research

A natural measure for the quality of NE solutions is the utilitarian social objectives. In this paper, we have considered such social objectives for two models, the classical load-balancing game and the cost-sharing game that is the same as the load-balancing one, except that there are an unlimited number of resources available but each comes with a set-up cost, which is proportionally shared by all users of the resource. The latter model is an extension of the former in the sense that both types of individual congestion costs are taken into account: positive and negative congestion effects, as we discussed at the end of the introduction section.

For the load-balancing game, we have identified a problem parameter and with this parameter we have provided sharp bounds for the PoS and PoA, which are asymptotically tight. Our work fills a gap in the literature on the well-studied load-balancing game.

For the cost-sharing game, we have used another problem parameter to provide a lower bound on the PoS and an upper bound on the PoA, which show that both the PoS and PoA are unbounded in general. Unfortunately, our results have left a gap: the two parametric bounds are not necessarily asymptotically tight. It is interesting to note that such a gap is very much similar (asymptotically) to the gap left in Berenbrink et al. (2006) for the load-balancing game as given in the introduction section, though our gap is much smaller.

An interesting question arises from the cost-sharing model. We have used a natural way of distributing an activation cost among the resource users with each user having a *proportional* share. An easy alternative is equal cost distribution among all the resource users, which we call *congestion sharing*. However, it is easy to see that the lower bound on the PoA demonstrated in Example 4.6 is still valid for *any* seemingly fair cost-sharing mechanism that charges the activation cost *only* to the users of that resource. In other words, no such *fair* mechanism can improve the game efficiency in terms of the PoA. However, improvement of equilibrium behavior could be achieved if we allow more freedom in distributing activation costs or even allow introduction of tolling mechanism. More specifically, as we have observed that inefficient equilibria are resulted in by infinitesimal jobs gathering on the same machine. A tolling mechanism to discourage such a behavior would reduce the PoA.

Another direction for further research is to investigate whether the two games we have considered are *smooth* as defined in Roughgarden (2009), so that our results on the PoA can be extended to alternative classes of equilibria.

References

Aland, S., Dumrauf, D., Gairing, M., Monien, B., & Schoppmann, F. (2006). Exact price of anarchy for polynomial congestion games. In B. Duran & W. Thomas (Eds.), *LNCS: Vol. 3884. Proceedings of the 23rd international symposium on theoretical aspects of computer science* (pp. 218–229). Berlin: Springer.

Anshelevich, E., Dasgupta, A., Kleinberg, J., Tardos, E., Wexler, T., & Roughgarden, T. (2008). The price of stability for network design with fair cost allocation. *SIAM Journal on Computing*, *38*(4), 1602–1623.

Awerbuch, B., Azar, Y., & Epstein, A. (2005). The price of routing unsplittable flow. In *37th ACM symposium on theory of computing* (pp. 57–66), Baltimore, MD, USA, 22–24 May 2005.

Berenbrink, P., Goldberg, L. A., Goldberg, P. W., & Martin, R. (2006). Utilitarian resource assignment. *Journal of Discrete Algorithms*, *4*, 567–587.

Czumaj, A. (2004). Selfish routing on the Internet. In *Handbook of scheduling: algorithms, models, and performance analysis* (pp. 517–542). Boca Raton: CRC Press, Chap. 42.

Czumaj, A., & Vöcking, B. (2002). Tight bounds for the worst-case equilibria. In *13th annual ACM-SIAM symposium on discrete algorithms* (pp. 413–420). Philadelphia: ACM-SIAM.

Endriss, U., Maudet, N., Sadri, F., & Toni, F. (2003). On optimal outcomes of negotiations over resources. In *The 2nd international joint conference on autonomous agents and multiagent systems (AAMAS)* (pp. 177–184), Melbourne, Australia, 2003.

Epstein, A., Feldman, M., & Mansour, Y. (2007). Strong equilibrium in cost sharing connection games. In *EC '07: Proceedings of the 8th ACM conference on electronic commerce* (pp. 84–92). New York: ACM.

Feigenbaum, J., Papadimitriou, C. H., & Shenker, S. (2001). Sharing the cost of multicast transmissions. *Journal of Computer and System Sciences*, *63*(1), 21–41.

Feldman, M., & Tamir, T. (2008). Conflicting congestion effects in resource allocation games. In G. Goos, J. Hartmanis, & J. van Leeuwen (Eds.), *LNCS: Vol. 5385*. *Proceedings of the 4th international workshop on Internet and network economics* (pp. 109–117. Berlin/Heidelberg: Springer.

Gairing, M., & Scoppmann, F. (2007). Total latency in singleton congestion games. In X. Deng & F. C. Graham (Eds.), *LNCS: Vol. 4858*. *Proceedings of the 3rd workshop on Internet and network economics (WINE 2007)* (pp. 381–387. Berlin/Heidelberg: Springer.

Gairing, M., Lücking, T., Mavronicolas, M., Monien, B., & Rode, M. (2008a). Nash equilibria in discrete routing games with convex latency functions. *Journal of Computer and System Sciences*, *74*, 1199–1225.

Gairing, M., Monien, B., & Tiemann, K. (2008b). Selfish routing with incomplete information. *Theory of Computing Systems*, *42*, 91–130.

Heydenreich, B., Müller, R., & Uetz, M. (2007). Games and mechanism design in machine scheduling—an introduction. *Production and Operations Management*, *16*(4), 437–454.

Hoefer, M., & Souza, A. (2011). Tradeoffs and average-case equilibria in selfish routing. *ACM Transactions on Computation* Theory (to appear).

Koutsoupias, E., & Papadimitriou, C. (1999). Worst-case equilibria. In G. Goos, J. Hartmanis, & J. van Leeuwen (Eds.), *LNCS: Vol. 1563*. *Proceedings of the 16th international symposium on theoretical aspects of computer science* (pp. 404–413. Berlin: Springer.

Lücking, T., Mavronicolas, M., Monien, B., & Rode, M. (2008). A new model for selfish routing. *Theor. Comput. Sci.*, *406*(3), 187–206.

McBurney, P., Parsons, S., & Wooldridge, M. (2002). Desiderata for argumentation protocols. In *The 1st international joint conference on autonomous agents and multiagent systems (AAMAS)* (pp. 402–409), Bologna, Italy.

Papadimitriou, C. (2001). Algorithms, games, and the Internet. In *Proc. 33rd ACM symposium on theory of computing* (pp. 749–753). New York: ACM Press.

Roughgarden, T. (2009). Intrinsic robustness of the price of anarchy. In *STOC '09: Proceedings of the 41st annual ACM symposium on theory of computing* (pp. 513–522). New York: ACM.

Roughgarden, T., & Tardos, É. (2002). How bad is selfish routing? *Journal of ACM*, *49*(2), 236–259.

Sandholm, T. W. (1998). Contract types for satisficing task allocation: I theoretical results. In *The AAAI spring symposium: satisficing models*, Palo Alto, California, USA.

Vöcking, B. (2007). Selfish load balancing. In *Algorithmic game theory* (pp. 517–542). Cambridge: Cambridge University Press.