Brief Announcement: Approximation Algorithms for Unsplittable Resource Allocation Problems with Diseconomies of Scale

Antje Bjelde*
School of Business and Economics
Humboldt University Berlin
Spandauer Straße 1
Berlin 10789, Germany
antje.bjelde@hu-berlin.de

Max Klimm†
School of Business and Economics
Humboldt University Berlin
Spandauer Straße 1
Berlin 10789, Germany
max.klimm@hu-berlin.de

Daniel Schmand
School of Business and Economics
RWTH Aachen University
Kackertstraße 7
Aachen 52072, Germany
schmand@oms.rwth-aachen.de

ABSTRACT

We study general resource allocation problems with a diseconomy of scale. Given a finite set of commodities that request certain resources, the cost of each resource grows superlinearly with the demand for it, and our goal is to minimize the total cost of the resources. In large systems with limited coordination, it is natural to consider local dynamics where in each step a single commodity switches its allocated resources whenever the new solution after the switch has smaller total cost over all commodities. This yields a deterministic and polynomial time algorithm with approximation factor arbitrarily close to the *locality gap*, i.e., the worst case ratio of the cost of a local optimal and a global optimal solution. For costs that are polynomials with non-negative coefficients and maximal degree d, we provide a locality gap for weighted problems that is tight for all values of d. For unweighted problems, the locality gap asymptotically matches the approximation guarantee of the currently best known centralized algorithm [Makarychev, Srividenko FOCS14] but only requires local knowledge of the commodities.

CCS CONCEPTS

• **Theory of computation → Routing and network design problems**; **Distributed algorithms**; **Network games**;

KEYWORDS

Approximation Algorithm; Distributed Algorithm; Locality Gap; Energy Efficiency; Congestion Games

ACM Reference format:
Antje Bjelde, Max Klimm, and Daniel Schmand. 2017. Brief Announcement: Approximation Algorithms for Unsplittable Resource Allocation Problems with Diseconomies of Scale. In *Proceedings of SPAA '17, Washington DC, USA, July 24-26, 2017,* 3 pages.
https://doi.org/http://dx.doi.org/10.1145/3087556.3087597

*Work of this author is supported by Einstein Center for Mathematics Berlin and Caroline von Humboldt Completion Grant.
†Work of this author is supported by Einstein Center for Mathematics Berlin.

ACM ISBN 978-1-4503-4593-4/17/07…$15.00
https://doi.org/http://dx.doi.org/10.1145/3087556.3087597

1 INTRODUCTION AND RELATED WORK

We consider a finite set of resources R and a finite set of commodities $N = \{1, \ldots, n\}$. For each commodity $i \in N$, we are given (explicitly or implicitly) a set $S_i \subseteq 2^R$ of feasible solutions and a weight $w_i \in \mathbb{R}_{\geq 0}$. The cost of each resource r is determined by a non-negative convex function $c_r : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$. Given a solution $\mathbf{s} = (s_1, \ldots, s_n)$ with $s_i \in S_i$, the cost for commodity i is defined as $C_i(\mathbf{s}) := w_i \sum_{r \in s_i} c_r(w_r(\mathbf{s}))$, where $w_r(\mathbf{s}) := \sum_{j \in N : r \in s_j} w_j$ is the total weight put on resource r under solution $\mathbf{s}$. We are interested in finding solutions $\mathbf{s}$ that minimize the total cost of the commodities, i.e., that minimize $C(\mathbf{s}) := \sum_{i=1}^n C_i(\mathbf{s}) = \sum_{r \in R} w_r(\mathbf{s}) c_r(w_r(\mathbf{s}))$. We denote the space of all feasible solutions by $\mathbf{S} = S_1 \times S_2 \times \cdots \times S_n$ and an instance of a resource allocation problem by $I = (\mathbf{S}, (w_i)_{i \in N}, (c_r)_{r \in R})$. A resource allocation problem is called *unweighted* if $w_i = 1$ for all $i \in N$ and *weighted*, otherwise. For a solution $\mathbf{s} \in \mathbf{S}$ and a commodity $i \in N$, we write $\mathbf{s} = (s_i, \mathbf{s}_{-i})$ implying that $s_i \in S_i$ and $\mathbf{s}_{-i} \in S_1 \times \cdots \times S_{i-1} \times S_{i+1} \times \cdots \times S_n$.

We call a solution *local optimal*, if its cost cannot be decreased by reallocating a single commodity, that is, $C(\mathbf{s}) \leq C(s_i', \mathbf{s}_{-i})$ for all $i \in N$ and $s_i' \in S_i$. We are interested in the cost guarantees achieved by local optimal solutions. These guarantees are captured by the notion of a locality gap. For an instance I of a resource allocation problem, the *locality gap* is defined as $\max\{C(\mathbf{s})/C(\mathbf{s}') : \mathbf{s}, \mathbf{s}' \in \mathbf{S}$ and $\mathbf{s}$ is a local optimum$\}$.

Minimization problems of this kind appear in different contexts. In the area of energy efficient algorithms, resources model computing devices that can run at different speeds. As the speed is increased, the energy consumption increases superlinearly. Frequently, it is assumed that energy consumption can be modeled by a function $C_r(x) = k_r x^{q_r} = x k_r x^{q_r - 1}$ of the speed x where $k_r > 0$ and $q_r \in (1, 3]$ are device-specific parameters, see, e.g., Albers [3] for a reference. In the area of traffic networks, resources correspond to roads in a street network. The time needed to traverse a road increases with the total traffic on the road. Popular models are the travel time functions put forward by the US Bureau of Public Roads (BPR) which are of the form $c_r(x) = k_r(1 + 0.15(x/z_r)^4)$ where $k_r > 0$ and $z_r > 0$ are road-specific parameters and x is the load of the resource, see [13]. When minimizing the total congestion cost, we are interested in minimizing the average travel time of a unit of demand which is given by $C_r(x) = x c_r(x)$.

As a running example, it is useful to consider the most natural case where the set of resources R corresponds to the set of edges E

of a graph $G = (V, E)$. Each commodity is specified by a source node u_i and a target v_i. The set S_i of feasible solutions of commodity i corresponds to the set of simple (u_i, v_i)-paths in G. We note, however, that all our results continue to hold in a more general setting where the set of feasible solutions $S_i \subset 2^R$ is arbitrary. We only need to make the minimal assumption that each commodity can efficiently optimize over its set S_i as long as the vector of feasible solution $\mathbf{s}_{-i}$ of the other commodities is fixed. In fact, we only need to require that this optimization over S_i can be done efficiently within arbitrary precision. We assume:

(1) *For every constant $\alpha \geq 1$, every commodity $i \in N$, there is a polynomial algorithm (oracle) $\mathfrak{O}_{i,\alpha} : \mathrm{S}_{-i} \to S_i$ that, given a partial solution $\mathbf{s}_{-i} \in \mathrm{S}_{-i}$ as input, computes a feasible solution $s_i' \in S_i$ with $C(s_i', \mathbf{s}_{-i}) \leq \alpha \min_{s_i \in S_i} C(s_i, \mathbf{s}_{-i})$.*
(2) *A feasible solution $\mathbf{s} \in \mathrm{S}$ can be computed in polynomial time.*

Clearly, when the sets S_i corresponds to the set of paths in a network, this assumption is satisfied as shortest paths can be computed efficiently.

Currently, the approximation algorithm with the best performance guarantee for this problem is due to Makarychev and Sviridenko [9]. They propose a convex programming relaxation of the problem and show that for monomials with degree d, the integrality gap is equal to the $(d+1)$-st Bell number B_{d+1}. Using a randomized rounding technique, this yields a randomized algorithm with approximation guarantee $B_{d+1} + \epsilon$ for any $\epsilon > 0$ which can be shown to be of order $[O\left(d/\log d\right)]^{d+1}$, see [5]. However, the linear programming approach of Makarychev and Sviridenko has the disadvantage that it requires a large amount of central coordination as the convex program has to be solved by a single central authority and the routing decisions of the commodities are based on the solution of the linear programming relaxation. This may be infeasible to implement in large systems without a central authority that is able to collect the data and solve the program. In fact, when optimizing large decentralized networks such as the Internet with respect to the energy consumption or the total delay, such a central authority is usually absent.

For the optimization of such decentralized systems, it is natural to consider improvement dynamics where the system starts in an arbitrary state $\mathbf{s}$ and at each step one commodity chooses another path if it decreases their share of the cost, i.e., commodity i switches from path P to path Q if

$$\sum_{r \in P \setminus Q} c_r(w_r(\mathbf{s})) > \sum_{r \in Q \setminus P} c_r(w_r(\mathbf{s}) + w_i),$$

where $w_r(\mathbf{s})$ denotes the load of resource r in state $\mathbf{s}$. The stable points of these dynamics are the Nash equilibria of the underlying congestion game. There is a vast amount of literature quantifying the price of anarchy of congestion games, i.e., determining the worst case ratio of the total cost in a Nash equilibrium and the total cost of an optimal solution. Most relevant to our work, Aland et al. [2] showed that the price of anarchy for both weighted and unweighted congestion games with polynomial travel time functions c_r with maximal degree d (that correspond to polynomial total cost functions C_r with maximum degree $d+1$) is of order $[O(d/\log d)]^{d+1}$. Unfortunately, this result does not yield an efficient distributed approximation algorithm since for weighted commodities, the improvement dynamics may cycle (cf. Harks et al. [8]) and Nash equilibria may even be fully absent (cf. Harks and Klimm [7]). For unweighted commodities the improvement dynamics are guaranteed to converge by a potential function argument due to Rosenthal [11], but the convergence may take a number of steps that is exponential in the size of the underlying network, see Ackermann et al. [1]. These issues make these dynamics unfavorable in practice.

In order to obtain polynomial convergence for the unweighted case, Awerbuch et al. [4] studied an approximate version of the improvement dynamics where each commodity only switches if the potential function drops by a factor of $1+\delta$. They showed that after a polynomial sequence of δ-best replies, a solution is reached which is an $(5/2 + O(\delta))$-approximation to the minimal congestion cost in the case of linear cost functions. For polynomial cost functions with maximum degree d and positive coefficients, their approach yields a $(d^{d-o(1)} + O(\delta))$-approximation. The technique of Awerbuch et al. however, relies on the existence of a potential function and thus, cannot be applied to resource allocation problems with weighted commodities and non-linear cost functions.

To overcome this issue, we study a different improvement dynamics where the system again starts in an arbitrary state $\mathbf{s}$, but all commodities take the impact of their path choices on the other commodities into account. More formally, we require that commodity i switches from path P to path Q if this switch decreases the overall cost of the solution, i.e.,

$$\sum_{r \in P \setminus Q} \Big(w_r(\mathbf{s}) c_r\big(w_r(\mathbf{s})\big) - \big(w_r(\mathbf{s}) - w_i\big) c_r\big(w_r(\mathbf{s}) - w_i\big) \Big)$$
$$> \sum_{r \in Q \setminus P} \Big(\big(w_r(\mathbf{s}) + w_i\big) c_r\big(w_r(\mathbf{s}) + w_i\big) - w_r(\mathbf{s}) c_r\big(w_r(\mathbf{s})\big) \Big).$$

This approach has the advantage that the total cost of the current solution is monotonically decreasing in each step, which implies that the dynamics reach a local optimum after a finite number of steps. In contrast to the long history of papers quantifying the efficiency of Nash equilibria, much less is known regarding the efficiency of local optimal solutions. This is the main issue addressed in this paper.

By analyzing the locality gap of the local optimal solutions, we provide the first non-trivial analysis of a local search algorithm that is guaranteed to converge both for the weighted and unweighted case of the problem.

2 OUR RESULTS AND TECHNIQUES

We study local improvement dynamics for general resource allocation problems with convex costs. In particular, we are interested in quantifying the locality gap. We assume that the cost functions $c_r(x)$ are polynomials with non-negative coefficients and maximal degree d, i.e. $c_r(x) = \sum_{j \in J} \alpha_j^r x^j$ with $J \subset [0, d]$, $\alpha_j^r \geq 0$ for all $j \in J$ and all $r \in R$. This imposes a total cost of $x c_r(x)$ on resource r under load x.[1]

[1]This is the standard form for the total cost in the congestion game literature. The speed scaling literature usually writes the total cost of a resource as x^q for some $q > 1$. Clearly, both forms are equivalent up to a shift in the exponent.

Our general algorithm is based on local search. The local search starts in an arbitrary solution $\mathbf{s} \in \mathbf{S}$. In each iteration, the oracles $\mathfrak{O}_{i,\alpha}$ are called to find a commodity that has an alternative solution $s_i' \in S_i$ such that $(1+\delta)C(s_i', \mathbf{s}_{-i}) \leq C(\mathbf{s})$ for some sufficiently small $\delta > 0$. When no such commodity is found, the algorithm stops and returns the solution.

Orlin et al. [10] showed that every local search problem admits a PTAS in the sense that an $(1+\delta)$-approximate local optimal solution can be computed in polynomial time via approximate local improvements steps. In order to determine how well local optimal solutions approximate the global minimum, we use the notion of smoothness used by Chen et al. [6] to analyze the price of anarchy of altruistic versions of unweighted congestion games with linear costs. It builds upon a similar notion of (λ, μ)-smoothness due to Roughgarden [12].

For unweighted resource allocation problems, we show that the locality gap is of order $[O(d/\log d)]^{d+1}$.

Theorem 2.1. *For every unweighted resource allocation problem with polynomial cost functions with non-negative coefficients and maximal degree d the following hold for some $\alpha \in O((2d/\log d)^{d+1})$:*

(1) *The locality gap is α.*
(2) *For any $\epsilon > 0$ there is a polynomial time $\alpha+\epsilon$-approximation algorithm for the minimization of the total costs.*

For concrete values of d, we can evaluate the locality gap explicitly.

Theorem 2.2. *For any $\epsilon > 0$ there is a polynomial time $\alpha_d + \epsilon$-approximation for every unweighted resource allocation problem with polynomial cost functions with maximal degree d and non-negative coefficients, where $\alpha_1 = 3$, $\alpha_2 = 13$, $\alpha_3 = 61$, $\alpha_4 = 391$, and $\alpha_5 = 2,157$.*

For the case of linear cost functions, it has been shown by Chen et al. [6] that the bound on the locality gap of 3 is tight. We provide a lower bound on the locality gap for general polynomials with maximum degree d.

Theorem 2.3. *There is an unweighted resource allocation problem with monomial costs with degree d and locality gap $\lfloor \frac{d}{\log(d+2)} - 1 \rfloor^{d+1}$.*

For weighted resource allocation problems, we obtain a locality gap for all values of d.

Theorem 2.4. *For every weighted resource allocation problem with polynomial costs with non-negative coefficients and maximal degree d, the following hold for $\alpha = 1/(\sqrt[d+1]{2}-1)^{d+1} \in O\Big(((d+1)/\log 2)^{d+1}\Big)$:*

(1) *The locality gap is at most α.*
(2) *For any $\epsilon > 0$, there is a polynomial time $\alpha+\epsilon$-approximation algorithm for the minimization of the total costs.*

Specifically, the locality gap is $3 + 2\sqrt{2} \approx 5.829$ for $d = 1$ and the locality gap is $15\sqrt[3]{2} + 12\sqrt[3]{4} + 19 \approx 56.948$ for $d = 2$.

This locality gap is tight.

Theorem 2.5. *There is a resource allocation problem with monomial costs with degree d and locality gap arbitrarily close to $\Big(\sqrt[d+1]{2} - 1\Big)^{-(d+1)} \in O\Big((\frac{d+1}{\log 2})^{d+1}\Big)$.*

Further, the problem is APX-hard.

Theorem 2.6. *For all $\alpha < 1.02$, there is no polynomial time α-approximation algorithm for unweighted resource allocation problems with linear cost functions, unless* P = NP.

We can thus conclude: by considering approximate improvement dynamics where commodities switch to another path only when the total cost is decreased by at least a factor of $1+\delta$, where $\delta > 0$ is arbitrary, the dynamics are guaranteed to converge to an approximate local optimal solution in a polynomial number of steps. We show that by sacrificing an additional factor of $1+\delta$, all locality gaps above yield a *deterministic*, *combinatorial* and *distributed* algorithm for the minimization of the total cost with the corresponding approximation guarantee. For weighted resource allocation problems, this yields the first deterministic algorithm, the first combinatorial algorithm, and the first distributed algorithm with non-trivial approximation guarantee as the algorithm of Makarychev and Srividenko relies on the centralized solution of a linear program and the randomized rounding of its solution.

For unweighted problems with linear cost, our asymptotic approximation guarantee of $[O(d/\log d)]^{d+1}$ is a slight improvement over the asymptotic behavior of the algorithm by Awerbuch et al. whose asymptotic behavior is of order $d^{d-o(1)}$. Our calculation of the concrete approximation guarantees for maximum degree up to ten suggests that our approximation is in fact better than the one by Awerbuch et al. for all $d \geq 9$. The approximation guarantee of $[O(d/\log d)]^{d+1}$ asymptotically matches the currently best known guarantee of Makarychev and Srividenko, but in contrast to the their algorithm, our algorithm is deterministic, can be implemented in a distributed fashion, and does not rely on the approximate solution of a convex program.

REFERENCES

[1] H. Ackermann, H. Röglin, and B. Vöcking. On the impact of combinatorial structure on congestion games. *J. ACM*, 55(6):1–22, 2008.
[2] S. Aland, D. Dumrauf, M. Gairing, B. Monien, and F. Schoppmann. Exact price of anarchy for polynomial congestion games. *SIAM J. Comput.*, 40(5):1211–1233, 2011.
[3] S. Albers. Energy-efficient algorithms. *Commun. ACM*, 53:86–96, 2010.
[4] B. Awerbuch, Y. Azar, A. Epstein, V. S. Mirrokni, and A. Skopalik. Fast convergence to nearly optimal solutions in potential games. In *Proc. 9th ACM Conf. Electron. Commerce*, pages 264–273, 2008.
[5] D. Berend and T. Tassa. Improved bounds on bell numbers and on moments of sums of random variables. *Probability and Mathematical Statistics*, 30:185–205, 2010.
[6] P.-A. Chen, B. De Keijzer, D. Kempe, and G. Schäfer. The robust price of anarchy of altruistic games. In N. Chen, E. Elkind, and E. Koutsoupias, editors, *Proc. 7th Int. Workshop Internet and Network Econ.*, volume 7090 of *LNCS*, pages 383–390, 2011.
[7] T. Harks and M. Klimm. On the existence of pure Nash equilibria in weighted congestion games. *Math. Oper. Res.*, 37(3):419–436, 2012.
[8] T. Harks, M. Klimm, and R. H. Möhring. Characterizing the existence of potential functions in weighted congestion games. *Theory Comput. Syst.*, 49(1):46–70, 2011.
[9] K. Makarychev and M. Sviridenko. Solving optimization problems with diseconomies of scale via decoupling. In *Proc. 55th Annu. IEEE Symp. Found. Comput. Sci.*, pages 571–580, 2014.
[10] J. B. Orlin, A. P. Punnen, and A. S. Schulz. Approximate local search in combinatorial optimization. *SIAM J. Comput.*, 33(5):1201–1214, 2004.
[11] R. W. Rosenthal. A class of games possessing pure-strategy Nash equilibria. *Internat. J. Game Theory*, 2(1):65–67, 1973.
[12] T. Roughgarden. Intrinsic robustness of the price of anarchy. In *Proc. 41st Annu. ACM Symp. Theory Computing*, pages 513–522, 2009.
[13] U.S. Bureau of Public Roads. *Traffic assignment manual.* U.S. Department of Commerce, Urban Planning Division, Washington, DC, 1964.