MHCDMC Rouding

Qinghui Zhang[1], Weidong Li[2], Qian Su[1] Xuejie Zhang[1, *]

1. School of Information Science and Engineering, Yunnan University, Kunming 650504, China

2. School of Mathematics and Statistics, Yunnan University, Kunming 650504, China

October 11, 2022

Abstract

abstract

keywords: high performance computing; resource allocation; scheduling; approximation algorithm; bag-of-tasks

*Correspondence: xjzhang@ynu.edu.cn (X. Zhang)

1 Introduction

introduction

2 Related work

related work

3 Online scheduling model

??? Assume that a heterogeneous computing system consists of m heterogeneous machines and n users. Each user i submit a set of a_i independent tasks known as a bag-of-tasks [2] and its profit p_i. As frequently used in scheduling problems for heterogeneous computing systems [23], let $\mathbf{ETC} = (ETC_{ij})$ be a $n \times m$ matrix where ETC_{ij} is the *estimated time to compute* for a task of user i on machine j. Similarly, let $\mathbf{APC} = (APC_{ij})$ be a $n \times m$ matrix where APC_{ij} is the *average power consumption* for a task of user i on machine j. Let x_{ij} be the number of tasks of user i assigned to machine j, where x_{ij} is the primary decision variable in the optimization problem. For a feasible solution $\mathbf{x} = (x_{ij})$, the load (or finishing time) of machine j is defined as

$$L_j = \sum_{i=1}^{n} ETC_{ij} x_{ij}. \tag{1}$$

It implies that the maximum finishing time of all machines (i.e., *makespan*), denoted by $MS(\mathbf{x})$, is

$$MS(\mathbf{x}) = \max_{j} L_j. \tag{2}$$

Correspondingly, the energy consumed by n users is given by:

Let c be the cost per unit of energy. Motivated by the offline model in [25], we consider the *Energy-Aware Profit Maximizing* (EAPM, for short) problem with bag-of-tasks, which can be formulated as the following nonlinear integer program (NLIP, for short).

$$Maximize_{\mathbf{x}} \quad \frac{\sum_{i=1}^{n} p_i - cE(\mathbf{x})}{MS(\mathbf{x})} \tag{3}$$

$$\text{subject to:} \quad \sum_{j=1}^{m} x_{ij} = a_i, \forall i = 1, 2, \ldots, n \tag{4}$$

$$\sum_{i=1}^{n} x_{ij} ETC_{ij} \leq MS(\mathbf{x}), \forall j = 1, 2, \ldots, m \tag{5}$$

$$x_{ij} \in \mathbb{Z}_{\geq 0}, \forall i, j. \tag{6}$$

The objective of (4) is to maximize the profit per unit time, where $\mathbf{x}$ is the primary decision variable. The first constraint ensures that each task in the bag is assigned to some machine. Because the objective is to maximize the profit per unit time, which is equivalent to minimize makespan, the second constrain ensures that $MS(\mathbf{x})$ is equal to the maximum finishing time of all machines.

However, in practice, when a user arrives, we have to assign all the tasks to machines as we do not know the information of the uncoming users. Thus, it is necessary to study the online EAPM problem with bag-of-tasks, where the tasks of user i have to assigned before the user $i+1$ arrives, for $i = 1, 2, \ldots, n-1$. Without loss of generality, assume each task of user i must be assigned to some machine before the tasks of user $i+1$ arrive, for $i = 1, \ldots, n-1$. Most importantly, the number of tasks of user i is very large, we can not assign the a_i tasks one by one. It motivates us to design an efficient algorithm for the online EAPM problem with bag-of-tasks.

4 An online algorithm

In this section, we present an efficient algorithm for the online EAPM problem with bag-of-tasks. For each i, let L_j^i and E^i be the *load* of machine j and the total energy consumed after assigning the tasks of the first i users. Initially, let $L_j^0 = 0$ for $j = 1, \ldots, M$ and $E^0 = 0$. By definitions, for $i = 1, 2, \ldots, n$, we have

$$L_j^i = \sum_{k=1}^{i} x_{kj} ETC_{kj}, \text{ and } E^i = \sum_{k=1}^{i} \sum_{j=1}^{m} x_{kj} APC_{kj} ETC_{kj}. \tag{7}$$

For $i = 1, 2, \ldots, n$, when user i arrives, we shall decide x_{ij} such that $\sum_{j=1}^{m} x_{ij} = a_i$ and the objective value

$$\frac{\sum_{k=1}^{i} p_k - cE^{i-1} - c\sum_{j=1}^{m} x_{ij} APC_{ij} ETC_{ij}}{MS^i} \tag{8}$$

is maximized, where

$$MS^i = \max_j L_j^i, \text{ and } L_j^i = L_j^{i-1} + x_{ij}ETC_{ij}, \forall i, j. \tag{9}$$

Formally, this problem can be formulated as the following integer program (IP):

$$\begin{cases} \text{Maximize } \dfrac{\sum_{k=1}^{i} p_k - cE^{i-1} - c\sum_{j=1}^{m} x_{ij}APC_{ij}ETC_{ij}}{MS^i} \\ \displaystyle\sum_{j=1}^{m} x_{ij} = a_i \\ L_j^{i-1} + x_{ij}ETC_{ij} \leq MS^i \\ x_{ij} \in Z^+ \cup \{0\}, j = 1, \ldots, m. \end{cases} \tag{10}$$

Note that inequality in IP(11) is equivalent to

$$x_{ij} \leq \lfloor \frac{MS^i - L_j^{i-1}}{ETC_{ij}} \rfloor. \tag{11}$$

For convenience, sort the tasks of user i in descending order by $APC_{ij}ETC_{ij}$. Without loss of generality, assume that

$$APC_{i1}ETC_{i1} \geq APC_{i2}ETC_{i2} \geq \cdots \geq APC_{im}ETC_{im}. \tag{12}$$

Our algorithm is based on the following lemma.

Lemma 1. *There exists an optimal solution such that*

$$x_{i1} = \cdots = x_{i(\tau-1)} = 0, \textit{and } x_{ij} = \lfloor \frac{MS^i - L_j^{i-1}}{ETC_{ij}} \rfloor, j = \tau, \ldots, m,$$

for some $\tau \in \{1, \ldots, M\}$.

Proof. Assume that we know the value of MS^i in the optimal solution for IP(11). Thus, the objective function of IP(11) is equivalent to minimize $\sum_{j=1}^{m} x_{ij}APC_{ij}ETC_{ij}$, as L_j^{i-1} and ETC_{ij} are constants. Obviously, to minimize $\sum_{j=1}^{m} x_{ij}APC_{ij}ETC_{ij}$, x_{ij} with small value $APC_{ij}ETC_{ij}$ should be maximized and x_{ij} with large value $APC_{ij}ETC_{ij}$ should be minimized, subject to the constraints of IP(11).

In the optimal solution $(x_{i1}, x_{i2}, \ldots, x_{im})$ for IP(11), consider the machine with minimum index τ_1 such that $x_{i\tau_1} > 0$. If there exists a machine τ_2 $(\geq \tau_1)$ such that $x_{i\tau_2} < \lfloor \frac{MS^i - L_{\tau_2}^{i-1}}{ETC_{i\tau_2}} \rfloor$, set

$$x'_{ij} = \begin{cases} x_{ij} - 1, \ \text{if } j = \tau_1 \\ x_{ij} + 1, \ \text{if } j = \tau_2 \\ x_{ij}, \ \text{if } j \neq \tau_1, \tau_2. \end{cases} \tag{13}$$

It is easy to verify that $(x'_{i1}, x'_{i2}, \ldots, x'_{im})$ is a feasible solution for IP(11) whose objective value is no less than that of $(x_{i1}, x_{i2}, \ldots, x_{im})$, as $APC_{i\tau_1} ETC_{i\tau_1} \geq APC_{i\tau_2} ETC_{i\tau_2}$ (see (13)). Repeat the above process, until that $x_{ij} = \lfloor \frac{MS^i - L_j^{i-1}}{ETC_{ij}} \rfloor$, for any machine j $(\geq \tau)$, where τ is the minimum machine index such that $x_{i\tau} > 0$. It implies that we find an optimal solution such that

$$x_{i1} = \cdots = x_{i(\tau-1)} = 0, \text{and } x_{ij} = \lfloor \frac{MS^i - L_j^{i-1}}{ETC_{ij}} \rfloor, \ \text{for } j = \tau + 1, \ldots, m.$$

Thus, the theorem holds. ∎

Given the value of MS^i in the optimal solution, for $j = m$, $m - 1$, …, 1, assign $\lfloor \frac{MS^i - L_j^{i-1}}{ETC_{ij}} \rfloor$ tasks to machine j until all tasks are assigned. According to Lemma 1, we find an optimal solution. Although we do not know the value of MS^i in the optimal solution, MS^i must be in $\{L_j^{i-1} + kETC_{ij} | k = 1, 2, \ldots, a_i, j = 1, 2, \ldots, m\}$. Trying all possible values of MS^i (at most $O(ma_i)$), we can find the optimal value of MS^i.

For each $\tau = 1, \ldots, m$, we only consider the variables x_{ij} $(j = \tau, \ldots, m)$, which implies that we schedule tasks of type i to machines of type j $(j = \tau, \ldots, M)$. Solving x_{ij} is equivalent to solve a solution to the following system of linear equations:

$$\begin{aligned} &\sum_{j=\tau}^{m} x_{ij} = a_i; \\ &x_{ij} = \frac{MS^i - L_j^{i-1}}{ETC_{ij}}, j = \tau, \ldots, M \quad . \end{aligned} \tag{14}$$

Note that there are $m - \tau + 2$ equations and $m - \tau + 2$ variables MS_i and x_{ij} $(j = \tau, \ldots, M)$. Thus, this system of linear equations can be solved in polynomial time. For each $\tau = 1, \ldots, m$, we obtain a feasible solution x_{ij}. Comparing the objective values of these m solution, we can find the best solution. Then, for $j = m$ to 1, we assign $\lceil x_{ij} \rceil$ tasks of type i to machines of type j, until all tasks are assigned. It is no hard to verify that the overall running time is polynomial in n and m.

Algorithm 2 shows the pseudo-code for the online algorithm.

ALGORITHM 2 Online assigning the tasks of each type to machines.

```
For $i = 1$ to $n$ do
    Relabel the indices of tasks such that $APC_{i1}ETC_{i1} \geq \ldots \geq APC_{iM}ETC_{iM}$;
    For $\tau = 1$ to $m$ do
        Solve 14 to find a solution $x_{ij}^{\tau}$;
        Comparing these $M$ solution to find the best solution $x_{ij}$ such that
        $\frac{p-cE^{i-1}-c\sum_{j=1}^{M} x_{ij}APC_{ij}ETC_{ij}}{MS^i}$ is maximized.
    End for
    For $j = M$ to 1 do
        Assign $\lceil x_{ij} \rceil$ tasks of type $i$ to machines of type $j$, until all tasks are assigned;
        Update the $L_j$ of type $j$;
    End for
End for
```

5 Experimental Results

6 Conclusion and future work

We feel that the local assignment algorithm in Section will find application in related areas.

Acknowledgement

The work is supported in part by the National Natural Science Foundation of China [Nos. 61662088, 61762091], the Program for Excellent Young Talents, Yunnan University, and IRT-STYN.

References

[1] S. Abdi, L. PourKarimi, M. Ahmadi, F. Zargari, “ Cost minimization for deadline-constrained bag-of-tasks applications in federated hybrid clouds, ” *Future Generation Computer Systems* , vol. 71, pp. 113-128, 2017.

[2] T. D. Braun, H. J. Siegel, N. Beck, L. L. Boloni, M. Maheswaran, A. I. Reuther, J. P. Robertson, M. D. Theys, B. Yao, D. Hensgen, and R. F. Freund, “ A comparison of eleven static heuristics for mapping a class of independent tasks onto heterogeneous distributed computing systems, ” *Journal of Parallel and Distributed Computing*, vol. 61, no. 6, pp. 810-837, 2001.

[3] J.P. Champati, B. Liang, “Semi-online algorithms for computational task offloading with communication delay, ” *IEEE Transactions on Parallel and Distributed Systems* , vol. 28, no. 4, pp. 1189 - 1201, 2017.

[4] M. Dayarathna, Y. Wen, and R. Fa, “ Data center energy consumption modeling: a survey, ” *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 732-794 2016.

[5] C. O. Diaz, J. E. Pecero, and P. Bouvry, “ Scalable, low complexity, and fast greedy scheduling heuristics for highly heterogeneous distributed computing systems,” *The Journal of Supercomputing*, vol. 67, no. 3, pp. 837-853, 2014.

[6] V. Ebrahimirad, M. Goudarzi, and A. Rajabi, “ Energy-aware scheduling for precedence-constrained parallel virtual machines in virtualized data centers, ” *Journal of Grid Computing*, vol. 13, no. 2, pp. 233-253, 2015.

[7] R. Friese, T. Brinks, C. Oliver, H. J. Siegel, and A. A. Maciejewski, “Analyzing the trade-offs between minimizing makespan and minimizing energy consumption in a heterogeneous resource allocation problem, ” in *The Second International Conference on Advanced Communications and Computation*, pp. 81-89, 2012.

[8] R. Friese, B. Khemka, A.A. Maciejewski, H.J. Siegel, G.A. Koenig, S. Powers, M. Hilton, J. Rambharos, G. Okonski, and S. W. Poole, “An analysis framework for investigating the trade-offs between system performance and energy consumption in a heterogeneous computing

environment", in *IEEE 27th International Parallel and Distributed Processing Symposium Workshops*, pp. 19-30, 2013.

[9] R. Friese, T. Brinks, C. Oliver, H.J. Siegel, A.A. Maciejewski, and S. Pasricha, "A machine-by-machine analysis of a bi-objective resource allocation problem", in *International Conference on Parallel and Distributed Processing Technologies and Applications*, pp. 3-9, 2013.

[10] R. Graham, "Bounds on multiprocessing timing anomalies, " *SIAM Journal on Applied Mathematics*, vol. 17, no. 2, pp. 416-429, 1969.

[11] A. Hameed, A. Khoshkbarforoushha, R. Ranjan, P. P.h Jayaraman, J. Kolodziej, P. Balaji, S. Zeadally, Q. M. Malluhi, N. Tziritas, A. Vishnu, S. U. Khan, and A. Zomaya, "A survey and taxonomy on energy efficient resource allocation techniques for cloud computing systems," *Computing*, vol. 98, pp. 751-774, 2016.

[12] C. H. Hsu, K. D. Slagter, S. C. Chen, and Y. C. Chung, "Optimizing energy consumption with task consolidation in clouds, " *Information Sciences*, vol. 258, no. 10, pp. 452-462, 2014.

[13] N. Karmarkar, "A new polynomial-time algorithm for linear programming, " in *Proceedings of the sixteenth annual ACM symposium on Theory of computing (STOC)*, pp. 302-311, 1984.

[14] T. Kaur, and I. Chana, "Energy efficiency techniques in cloud computing: a survey and taxonomy, " *ACM Computing Surveys*, vol. 48, no. 2, Article No. 22, 2015.

[15] B. Khemka, R. Friese, S. Pasricha, A. A. Maciejewski, H. J. Siegel, G. A. Koenig, S. Powers, M. Hilton, R. Rambharos, and S. Poole, "Utility maximizing dynamic resource management in an over subscribed energy-constrained heterogeneous computing system," *Sustainable Computing: Informatics and Systems*, vol. 5, pp. 14-30, 2015.

[16] B. Khemka, R. Friese, L. D. Briceno, H. J. Siegel, A. A. Maciejewski, G. A. Koenig, C. Groer, G. Okonski, M. M. Hilton, R. Rambharos, and S. Poole, " Utility functions and resource management in an oversubscribed heterogeneous computing environment, " *IEEE Transactions on Computers*, vol. 64, no. 1, pp. 2394-2407, 2015.

[17] F. Juarez, J. Ejarque, and R. M. Badia, "Dynamic energy-aware scheduling for parallel task-based application in cloud computing," Future Generation Computer Systems, vol. 78, pp. 257-271, 2018.

[18] W. Li, X. Liu, X. Cai, X. Zhang, " Approximation algorithm for the energy-aware profit maximizing problem in heterogeneous computing systems," *Journal of Parallel and Distributed Computing*, vol. 124, 70-77, 2019.

[19] M. Sajid, Z. Raza, " Energy-aware stochastic scheduler for batch of precedence-constrained jobs on heterogeneous computing system," *Energy*, vol. 125, pp. 258-274, 2017.

[20] L. Mashayekhy, M.M. Nejad, D. Grosu, Q. Zhang, and W. Shi, " Energy-aware scheduling of mapreduce jobs for big data applications, " *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 10, pp. 2720-2733, 2015.

[21] A.-C. Orgerie, M. D. de Assuncao, and L. Lefevre, "A survey on techniques for improving the energy efficiency of large-scale distributed systems, " *ACM Computing Surveys*, vol. 46, no. 4, Article No. 47, 2014.

[22] M.A. Oxley, S. Pasricha, A.A. Maciejewski, H.J. Siegel, J. Apodaca, D. Young, L. Briceno, J. Smith, S. Bahirat, B. Khemka, A. Ramirez, and Y. Zou, " Makespan and energy robust stochastic static resource allocation of bags-of-tasks to a heterogeneous computing system, " *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 10, pp. 2791-2805, 2015.

[23] K. M. Tarplee, R. Friese, A. A. Maciejewski, and H. J. Siegel, "Scalable linear programming based resource allocation for makespan minimization in heterogeneous computing systems, " *Journal of Parallel and Distributed Computing*, vol. 84, pp. 76-86, 2015.

[24] K. M. Tarplee, R. Friese, A. A. Maciejewski, H. J. Siegel, and E. Chong," Energy and makespan tradeoffs in heterogeneous computing systems using efficient linear programming techniques, " *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 6, pp. 1633-1646, 2016.

[25] K.M. Tarplee, A.A. Maciejewski, and H.J. Siegel, "Energy-aware profit maximizing scheduling algorithm for heterogeneous computing systems, " in *14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, pp. 595-603, 2014.

[26] A. Tchana, N. D. Palma, I. Safieddine, and D. Hagimont, "Software consolidation as an efficient energy and cost saving solution," *Future Generation Computer Systems*, vol. 58, pp. 1-12, 2016.

[27] J. Zhang, N. Xie, X. Zhang, W. Li, "An online auction mechanism for cloud computing resource allocation and pricing based on user evaluation and cost," *Future Generation Computer Systems* , vol. 89, pp. 286-299, 2018.

[28] F. Zhang, J. Cao, K. Li, S.U. Khan, K. Hwang, "Multi-objective scheduling of many tasks in cloud platforms," *Future Generation Computer Systems* , vol. 37, pp. 309-320, 2014.

[29] W. Zheng, Y. Qin, E. Bugingo, D. Zhang, J. Chen, "Cost optimization for deadline-aware scheduling of big-data processing jobs on clouds," *Future Generation Computer Systems* , vol. 82, pp. 244-255, 2018.

[30] (2013, March) Coin-or clp. [Online]. Available: https://projects.coinor. org/Clp

[31] (2013, May) Intel core i7 3770k power consumption, thermal. [Online]. Available: http://openbenchmarking.org/result/ 1204229-SU-CPUMONITO81