

通信遅延を考慮した完全 k 分木の 近似タスクスケジューリングアルゴリズム

藤本典幸 萩原兼一
大阪大学 大学院情報科学研究科

タスクグラフのトポロジが木で、タスク処理時間と通信遅延がすべて単位時間である場合に限定しても、タスクスケジューリング問題は NP 困難であることが知られている。本論文では、タスクグラフのトポロジが完全 k 分木で、タスク処理時間がすべて単位時間であり、かつ、通信遅延が高々 $(k-1)$ 単位時間で一定である場合のタスクスケジューリング問題に対する精度保証が高々3の近似アルゴリズムを示す。特に利用できるプロセッサ数が k のべきの場合に対して精度保証が高々2の近似アルゴリズムを示す。

An Approximation Algorithm for Task Scheduling of a Complete K-ary Tree with Communication Delays

Noriyuki Fujimoto Kenichi Hagihara
Graduate School of Information Science and Technology, Osaka University

It is known that task scheduling problem of an intree with unit time tasks and unit time communication delays onto a limited number of processors is NP-hard. In this paper, we show that such a problem is approximable within a factor of 3 if we assume a complete k -ary intree with at most $(k-1)$ constant communication delay. Also we prove that the problem is approximable within a factor of 2 if a given number p of available processors is a power of k .

1 はじめに

本論文では、すべてのタスク処理時間が単位時間であり、通信遅延が一定で、かつ、トポロジが木であるタスクグラフを p 台のプロセッサにスケジューリングするタスクスケジューリング問題を考える。ここで p は入力で指定される正整数である。すべての通信遅延が単位時間である場合に限定しても、この問題は NP 困難であることが知られている [12]。

以降では、すべてのタスク処理時間が単位時間で、かつ、高さが h である完全 k 分木を $T_{k,h}$ と書く。本論文では、まず最初に、 $T_{k,h}$ を p 台のプロセッサにタスク数 n の線形時間でスケジューリングするアルゴリズム TOSH (Two Optimal phases Scheduling Heuristic) を提案する。次に本論文では、通信遅延が高々 $(k-1)$ で一定の場合、 $kh+1$ が $T_{k,h}$ のスケジュールのメイクスパンの下界であることを示す。 n/p は $T_{k,h}$ のスケジュールのメイクスパンの自明な下界である。本論文では、TOSH が出力するスケジュールのメイクスパンをこれらの下界と比較することにより、TOSH は最適メイクスパンの高々3倍のメイクスパンをもつスケジュー

ルを出力することを示す。特に p が k のべきのとき、TOSH は最適メイクスパンの高々2倍のメイクスパンをもつスケジュールを出力することを示す。

2 諸定義

並列計算をモデル化した重み付きの DAG (Directed Acyclic Graph) をタスクグラフ [5] と呼ぶ。タスクグラフの節点は並列計算のタスクを表す。節点 u の重み $\lambda(u)$ は、タスク u の処理時間が $\lambda(u)$ 単位時間であることを意味する。有向辺 (u, v) は、タスク v の計算のためにはタスク u の計算結果が必要であることを意味する。 (u, v) の重み $\tau(u, v)$ は、タスク u を計算するプロセッサ q からタスク v を計算するプロセッサ q' へのプロセッサ間通信遅延は、 q と q' が異なる場合、高々 $\tau(u, v)$ 単位時間であることを意味する。 q と q' が同一なら、通信遅延は 0 とする。各タスクをいつどのプロセッサで実行するかを決定したものをスケジュールという。スケジュール S のすべてのタスクの実行が完了する時刻を S のメイクスパンという。メイクスパン最小のスケジュールを最適スケ

ジュールという．タスクグラフと利用可能なプロセッサ数が与えられて最適スケジュールを1つ見つける問題をタスクスケジューリング問題 (SPC) という．通信遅延が一定値 c をとる場合の SPC は以下の2つのクラスに分類できる [9] : G の節点数が増えると c の値も大きくなる SPC を一様通信遅延 スケジューリング問題 (UDS) という ; G に関わらず c が一定の SPC を定数通信遅延 スケジューリング問題 (CDS) という．最適メイクスパンの高々 ρ 倍のメイクスパンをもつスケジュールを ρ -近似スケジュールという． ρ -近似スケジュールを多項式時間で見つけるアルゴリズムを ρ -近似アルゴリズムという． ρ を近似率あるいは精度保証という [1] . 節点 u から節点 v に有向辺があるとき , u を v の直接先行節点という . 木の節点から根節点へのパス上の有向辺の数をその節点の深さという . 節点の深さの最大値を木の高さという . すべての内部節点がちょうど k 個の直接先行節点をもち , かつ , すべての葉の深さが等しい木を完全 k 分木という .

3 関連研究

本論文では , 問題の記述に Graham らの記法 [7] の拡張記法 [12] をさらに拡張した記法を用いる . 以下 , 拡張点について説明する . 記法 ' $2 \leq c_{jk} \leq 5$ ' は , 各通信遅延は 2 単位時間以上 5 単位時間以下であることを表す (通信遅延は一定でなくともよい) . 記法 ' $\text{general } c_{jk}$ ' , ' $\text{uniform } c_{jk}$ ' , ' $\text{constant } c_{jk}$ ' はそれぞれ問題が SPC , UDS , CDS であることを表す . 本論文の対象問題は , この記述法を用いると $\mathcal{P}|\text{complete } k\text{-ary intree}, p_j = 1, c_{jk} = c \ (1 \leq c \leq k-1)|C_{max}$ と書ける .

プロセッサ数が限られる場合については以下の結果が知られている . $\mathcal{P}|\text{intree}, p_j = 1, c_{jk} = 1|C_{max}$ は NP 困難である [12] . これに対して以下の制限を加えた問題は多項式時間で解けることが知られている : $\mathcal{P}_m|\text{intree}, p_j = 1, c_{jk} = 1|C_{max}$ は $O(n^{2m})$ 時間で解ける [18] ; $\mathcal{P}_2|\text{intree}, p_j = 1, c_{jk} = 1|C_{max}$ は $O(n)$ 時間で解ける [8, 11, 12, 15, 18] ; $\mathcal{P}_2$ with integer speed ratios |complete intree, $p_j = 1, c_{jk} = 1|C_{max}$ は $O(n)$ 時間で解ける [2] . 我々が知る限り , プロセッサ数が限られる場合の近似アルゴリズムは1つしか知られていない :

$\mathcal{P}|\text{prec}, c_{jk} = 1, p_j = 1|C_{max}$ の 3-近似アルゴリズム [16] が知られている .

プロセッサ数に制限がない場合については以下の結果が知られている . 木のスケジューリングの NP 困難性について以下の結果が知られている : $\mathcal{P}_\infty|\text{complete binary intree}, p_j = 1, \text{general } c_{jk}|C_{max}$ は NP 困難である [9] ; $\mathcal{P}_\infty|\text{binary intree}, p_j = 1, \text{uniform } c_{jk}|C_{max}$ は NP 困難である [9] ; $\mathcal{P}_\infty|\text{intree of height } \leq 2, \text{general } c_{jk}|C_{max}$ は NP 困難である [4] . これに対して以下の制限を加えた問題は多項式時間で解けることが知られている : $\mathcal{P}_\infty|\text{prec}, \text{constant } c_{jk}|C_{max}$ は $O(n^{\tau+1})$ 時間で解ける [10] . ここで τ は通信遅延である ; $\mathcal{P}_\infty|\text{complete } k\text{-ary intree}, \text{uniform } c_{jk}|C_{max}$ は $O(n^2 \log n)$ 時間で解ける [9] ; $\mathcal{P}_\infty|\text{intree}, c_{jk} \leq \min_j p_j|C_{max}$ は $O(n)$ 時間で解ける [3] ; $\mathcal{P}_\infty|\text{complete } k\text{-ary intree}, p_j = 1, c_{jk} = c \ (1 \leq c \leq k-1)|C_{max}$ は $O(n)$ 時間で解ける [6] ; $\mathcal{P}_\infty|\text{complete } k\text{-ary intree}, p_j = 1, k-1 \leq c_{jk} \leq k|C_{max}$ は $O(n)$ 時間で解ける [6] . 木のスケジューリングの近似アルゴリズムについては以下の結果が知られている : $\mathcal{P}_\infty|\text{prec}, p_j = 1, \text{uniform } c_{jk}|C_{max}$ に対する 2-近似アルゴリズム [14] が知られている ; $\mathcal{P}_\infty|\text{prec}, p_j = 1, c_{jk} = 1|C_{max}$ に対する $\frac{4}{3}$ -近似アルゴリズム [13] が知られている .

4 プロセッサ数が制限されない場合の最適アルゴリズム

TOSH がサブルーチンとして用いる Naive アルゴリズム [6] と , その性質について必要最小限の説明を行う .

アルゴリズム Naive

入力: すべての通信遅延が一定値 c である $T_{k,h}$

出力: プロセッサ数に制限がない場合の $T_{k,h}$ のスケジュール

begin

必要なだけプロセッサを用いて , すべての葉タスクを時刻 0 に割り当てる .

各深さ $i \ (0 \leq i \leq h-1)$ に対して以下を行う : 必要なだけプロセッサを用いて , 深さ i のすべてのタスクを時刻 $(c+1)(h-i)$ に割り当てる .

end

以下の定理が成り立つ [6] .

定理 1 すべての通信遅延が一定値 c である $T_{k,h}$ に対して Naive アルゴリズムはメイクスパン $(c+1)h+1$ のスケジュールを出力する。

定理 2 プロセッサ数に制限がない場合, $c \leq k-1$ ならば, $(c+1)h+1$ はすべての通信遅延が一定値 c である $T_{k,h}$ のスケジュールのメイクスパンの下界である。

5 提案アルゴリズム TOSH

k^h 個以上のプロセッサが利用できる場合は 4 節の Naive アルゴリズムにより $T_{k,h}$ の最適スケジュールが得られるので, 以下では一般性を失うことなく利用できるプロセッサ数は k^h 個未満であると仮定する。

アルゴリズム TOSH

入力: すべての通信遅延が一定値 c である $T_{k,h}$, 利用可能プロセッサ数 p

出力: $T_{k,h}$ の p プロセッサ用スケジュール

begin

r_i を深さ $\lceil \log_k p \rceil$ の節点とする ($1 \leq i \leq k^{\lceil \log_k p \rceil}$). T_i を r_i を根とする $T_{k,h}$ の部分木とする. T_0 を深さが高々 $(\lceil \log_k p \rceil - 1)$ の節点すべてからなる $T_{k,h}$ の部分木とする (図 1 参照).

P_i が $\{T_x | x = i + (j-1)p, 1 \leq j \leq \lceil k^{\lceil \log_k p \rceil} / p \rceil, 1 \leq x \leq k^{\lceil \log_k p \rceil}\}$ 中の全タスクをトポロジカル順に逐次計算するように割り当てる (図 1 参照).

t を r_i の実行完了時刻の最大値とする. 時刻 $t+c$ から T_0 を 4 節の Naive アルゴリズムを用いて p 個のプロセッサにスケジューリングする (図 1 参照).

end

以下の定理が成り立つ。

定理 3 $\lceil \log_k p \rceil - \log_k p$ を α と書く. $\forall (u, v) \in E, \tau(u, v) = c$ を満たす $T_{k,h}$ から TOSH は p 個のプロセッサに対してメイクスパン $(c+1)\lceil \log_k p \rceil + \frac{\lceil k^\alpha \rceil}{k^\alpha} \times \frac{k^{h+1}-pk^\alpha}{(k-1)p}$ のスケジュールを出力する。

6 TOSH の精度保証の解析

プロセッサ数に制限がない場合のメイクスパンの下界は, プロセッサ数が制限される場合の下界でもあるので, 定理 2 より, 以下の定理が成り立つ。

定理 4 p 個のプロセッサが利用できるとき, $c \leq k-1$ ならば, $(c+1)h+1$ はすべての通信遅延が一定値 c である $T_{k,h}$ のスケジュールのメイクスパンの下界である。

定理 5 利用できるプロセッサ台数が k のべきでない場合, TOSH は 3-近似アルゴリズムである。

証明 p 個のプロセッサが利用できるときの $T_{k,h}$ のスケジュールのメイクスパンの下界を $OPT_{k,h,p}$ とする. $T_{k,h}$ の総タスク数は $\frac{k^{h+1}-1}{k-1}$ である. ゆえに $OPT_{k,h,p} \geq \frac{k^{h+1}-1}{(k-1)p}$. また定理 4 より $OPT_{k,h,p} \geq (c+1)h+1$. 定理 3 より, TOSH はメイクスパン $(c+1)\lceil \log_k p \rceil + \frac{\lceil k^\alpha \rceil}{k^\alpha} \times \frac{k^{h+1}-pk^\alpha}{(k-1)p}$ のスケジュールを出力する. また $\alpha \geq 0$ かつ $k \geq 2$ より $\frac{\lceil k^\alpha \rceil}{k^\alpha} < 2$. ゆえに $\alpha \geq 0$ かつ $k \geq 2$ より $(c+1)\lceil \log_k p \rceil + \frac{\lceil k^\alpha \rceil}{k^\alpha} \times \frac{k^{h+1}-pk^\alpha}{(k-1)p} < ((c+1)h+1) + \frac{\lceil k^\alpha \rceil}{k^\alpha} \times \frac{k^{h+1}-1}{(k-1)p} \leq OPT_{k,h,p} + \frac{\lceil k^\alpha \rceil}{k^\alpha} \times OPT_{k,h,p} < 3OPT_{k,h,p}$ ■

系 1 k のべきの個数のプロセッサが利用できるとき, TOSH は 2-近似アルゴリズムである。

証明 p が k のべきのとき, 定理 5 の証明で $\alpha = 0$ である. ゆえに $\frac{\lceil k^\alpha \rceil}{k^\alpha} = 1$ より題意成立. ■

7 まとめ

すべてのタスク処理時間が単位時間で, かつ, 通信遅延が高々 $(k-1)$ 単位時間で一定である完全 k 分木のタスクスケジューリング問題に対する 3-近似アルゴリズムを示した. 特に p が k のべきである場合に対して 2-近似アルゴリズムを示した。

謝辞

本研究は一部, 日本学術振興会未来開拓学術研究推進事業 (JSPS-RFTF99I00903) および 栢森情報科学振興財団の補助による。

参考文献

- [1] 浅野孝夫: “近似アルゴリズム”, アルゴリズム工学: 計算困難問題への挑戦, 共立出版, pp.67-73, 2001
- [2] Blazewicz, J., Guinand, F., Penz, B., and Trystram, D.: “Scheduling Complete Trees on Two Uniform Processors with Integer Speed Ratios and Communication Delays”, Parallel Processing Letters, Vol.10, No.4, pp.267-277, 2000
- [3] Chrétienne, P.: “A Polynomial Algorithm to Optimally Schedule Tasks on a Virtual Distributed System under Tree-like Precedence Constraints”, European J. Oper. Res. Vol.43, pp.225-230, 1989

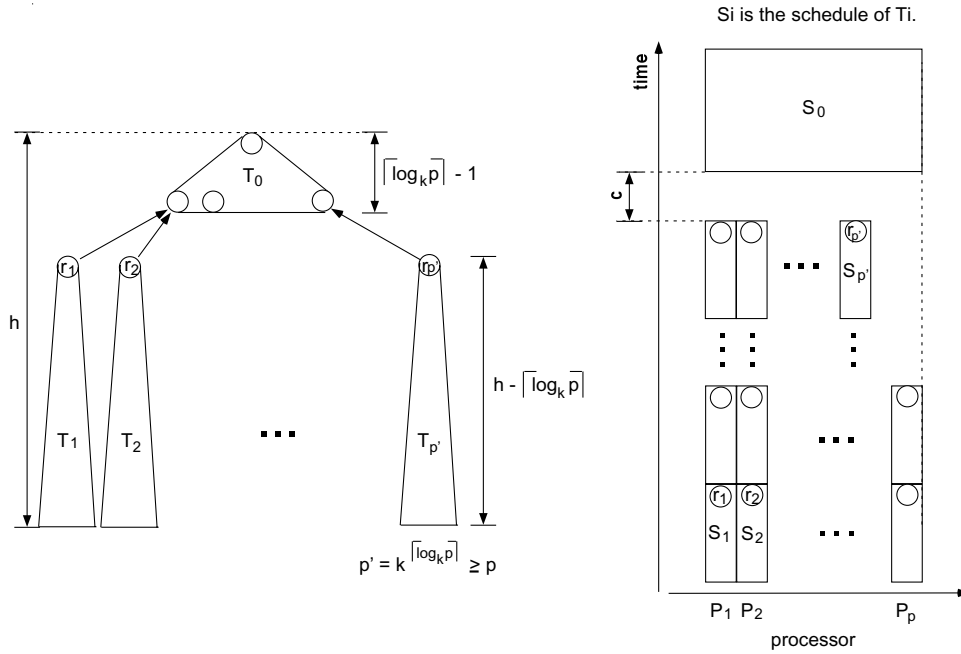


図 1: TOSH によるスケジューリング

- [4] Chrétienne, P. and Picouleau, C.: “Scheduling with Communication Delays: A Survey”, *Scheduling Theory and Its Applications*, pp.65–90, Wiley, 1995
- [5] El-Rewini, H., Lewis, T.G., and Ali, H.H.: “TASK SCHEDULING in PARALLEL and DISTRIBUTED SYSTEMS”, pp.56–105, PTR Prentice Hall, 1994.
- [6] Fujimoto, N. and Hagihara, K.: “Optimal Task Scheduling of a Complete K-ary Tree with Communication Delays”, 4th Int. Conf. on Parallel Processing and Applied Mathematics (PPAM2001), LNCS Vol.2328, to appear, 2002
- [7] Graham, R. L., Lawler, E. L., Lenstra, J. K., and Rinnooy Kan, A. H. G.: “Optimization and Approximation in Deterministic Sequencing and Scheduling: A Survey”, *Ann. Discrete Math.* Vol.5, pp.287–326, 1979
- [8] Guinand, F. and Trystram, D. : “Optimal Scheduling of UECT Trees on Two Processors”, *RAIRO Oper. Res.*, Vol.34, No.2, pp.131–144, 2000
- [9] Jakoby, A. and Reischuk, R.: “The Complexity of Scheduling Problems with Communication Delays for Trees”, 3rd Scandinavian Workshop on Algorithm Theory (SWAT’92), LNCS Vol.621, pp.165–177, 1992
- [10] Jung, H., Kirousis, L., and Spirakis, P.: “Lower bounds and Efficient Algorithms for Multiprocessor Scheduling of Dags with Communication Delays”, *the Information and Computation Journal*, Vol.105, No.1 pp.94–104, 1993
- [11] Lawler, E. L.: “Scheduling Trees on Multiprocessors with Unit Communication Delays”, Presented at the First Workshop on Models and Algorithms for Planning and Scheduling Problems, Villa Vigoni, Lake Como, Italy, unpublished manuscript, June, 1993
- [12] Lenstra, J. K., Veldhorst, M., and Veltman, B.: “The Complexity of Scheduling Trees with Communication Delays”, *Journal of Algorithms*, Vol.20, pp.157–173, 1996
- [13] Munier, A. and König, J.: “A Heuristic for a Scheduling Problem with Communication Delays”, *Oper. Res.*, Vol.45, No.1, pp.145–147, 1997
- [14] Papadimitriou, C.H. and Yannakakis, M.: “Towards an Architecture-Independent Analysis of Parallel Algorithms”, *SIAM J. Comput.* Vol.19, No.2, pp.322–328, 1990
- [15] Picouleau, C.: “Etude de Problèmes les Systèmes Distribués”, Ph.D. thesis, Univ. Pierre et Marie Curie, Paris, France, 1992
- [16] Rayward-Smith, V. J.: “Net Scheduling with Unit Interprocessor Communication Delays”, *Discrete Applied Mathematics*, Vol.18, pp.55–71, 1987
- [17] Thurimella, R. and Yesha, Y. : “A scheduling principle for precedence graphs with communication delay”, *Int. Conf. on Parallel Processing*, Vol.3, pp.229–236, 1992
- [18] Varvarigou, T., Roychowdhury, V. P., Kailath, T., and Lawler, E.: “Scheduling In and Out Forests in the Presence of Communication Delays”, *IEEE Trans. on Parallel and Distributed Systems*, Vol.7, No.10, pp.1065–1074, 1996