Bounds and approximation results for scheduling coupled-tasks with compatibility constraints
Rodolphe Giroudeau, Jean-Claude König, Benoit Darties, Gilles Simonin

To cite this version:

HAL Id: lirmm-01313722
https://hal-lirmm.ccsd.cnrs.fr/lirmm-01313722
Submitted on 7 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Bounds and approximation results for scheduling coupled-tasks with compatibility constraints

R. Giroudeau¹, J.C König¹, B. Darties² and G. Simonin³

¹ LIRMM UMR 5506, 161 rue Ada 34392, Montpellier France
² LE2I UMR6306, Univ. Bourgogne Franche-Comté, F-21000 Dijon, France
³ Insight Centre for Data Analytics, University College Cork, Ireland

Abstract. This article is devoted to propose some lower and upper bounds for the coupled-tasks scheduling problem in presence of compatibility constraints according to classical complexity hypothesis (\(P \neq NP\), ETH). Moreover, we develop an efficient polynomial-time approximation algorithm for the specific case for which the topology describing the compatibility constraints is a quasi-split-graph.

Keywords: coupled-task, compatibility graph, complexity, approximation.

1 Introduction, motivations, model

We consider in this paper the coupled-task scheduling problem subject to compatibility constraints. The motivation of this model is related to data acquisition processes using radar sensors: a sensor emits a radio pulse (first sub-task \(a_i\)), and listen for an echo reply (second sub-task \(b_i\)). To make the notation less cluttered, the processing time of a sub-task will be denoted by \(a_i\) instead of \(p_{a_i}\) used in the theory of scheduling. Between these two instants (emission and reception), clearly there is an idle time \(L_i\) due to the propagation, in both sides, of the radio pulse. A coupled-task \((a_i, L_i, b_i)\), introduced by Shapiro (1980), is a natural way to model such data acquisition. This model has been widely studied in several works, i.e. Blażewicz et al. (2009). Other works proposed a generalization of this model by including compatibility constraints: scheduling a sub-task during the idle time of another requires that both tasks are compatible. The relations of compatibility are modeled by a compatibility graph \(G\), linking pair of compatible tasks only. This model is detailed in Simonin et al. (2012). In previous works, we studied the complexity of scheduling coupled-tasks with compatible constraints under several parameters like the size of the sub-tasks or the class of the compatibility graph (Simonin et al. 2013).

In this work, we propose original complexity and approximation results for the problem of scheduling stretched coupled-task with compatibility constraints. A stretched coupled-tasks \(i\) is a coupled-task having both sub-tasks processing time and idle time equal to a triplet \((\alpha(i), \alpha(i), \alpha(i))\), where \(\alpha(i)\) is the stretch factor of the task \(i\) - one can apply a stretch factor \(\alpha(i)\) to a reference task \((1,1,1)\) to obtain \(i\).

The objective is to minimize the makespan \(C_{max}\). The input of the problem is a collection of coupled-tasks \(T = \{t_1, t_2, \ldots, t_n\}\), a stretch factor function \(\alpha : T \to \mathbb{N}\), and a compatibility graph \(G_c = (T, E)\) where edge from \(E\) link pairs of compatible tasks only. When dealing with stretched coupled-tasks only, a edge \(\{x, y\} \in E\) exists if \(\alpha(x) = \alpha(y)\) (then \(x\) and \(y\) can be scheduled together without idle time as the idle time of one task is employed to schedule the sub-task of the other, thus we can schedule sequentially \(a_x, a_y, b_x, b_y\).
The reduction is constructed in polynomial time.

in $G_X$ as follows: we define a set of tasks $G_X$ into $T$ is to find a matching (set of mutually disjoint triplets) $\text{Max-3DM-2}$ exponential-time hypothesis (see (Woeginger 2001) for a survey on exact algorithms for time algorithms under the following complexity-theoretic hypothesis that is known as the $\text{Max-3DM-2}$ $\text{NP}$-complete (Chlebik 2003). In this restricted version, we have $\text{Max-3DM-2}$ $\text{NP}$-complete on a quasi-split graph $G = (G_X, G_Y, E)$ even if $\#(V(G_X)) = 1$ and $\#(V(G_Y)) = 1$, but is 5/4-approximable.

2 Complexity and approximation results

Theorem 1. Deciding whether an instance of $1|\alpha, G_c = 1 - \text{stage - bipartite}, \#(X) = 2, \#(Y) = 1, d_{G_c}(X) \in \{1,2\}, d_{G_c}(Y) \in \{3,4\}|C_{\max}$ is a problem hard to approximate within $\frac{21}{22} - \rho^{\text{Max-3DM-2}} \leq \rho$, where $\rho^{\text{Max-3DM}}$ gives the upper bound for the $\text{Max-3DM}$.

Proof. We prove first that the problem is $\text{NP}$-complete via a polynomial-time reduction. Based on this reduction, we apply the gap-preserving reduction.

The proof is based on a reduction from the maximum 3 dimensional matching (MAX-3DM) (Garey & Johnson 1979): let $A, B,$ and $C$ be three disjoint sets of equal size, with $n = |A| = |B| = |C|$, and a set $T \subseteq A \times B \times C$ of triplet, with $|T| = m$. The aim is to find a matching (set of mutually disjoint triplets) $T^* \subseteq T$ of maximum size. This problem is well known to be $\text{NP}$-complete. The restricted version of this problem in which each element of $A \cup B \cup C$ appears exactly twice is denoted $\text{Max-3DM-2}$ and remains $\text{NP}$-complete (Chlebik 2003). In this restricted version, we have $m = 2n$.

We transform the instance of $\text{Max-3DM-2}$ to an instance of $1|\alpha, G_c = 1 - \text{stage bipartite}, \#(X) = 2, \#(Y) = 1, d_{G_c}(X) \in \{1,2\}, d_{G_c}(Y) \in \{3,4\}|C_{\max} = 63n - 3k(1 - \epsilon)$ as follows: we define a set of tasks $X \cup Y$ and model the compatibility constraint with a graph $G_c = (X, Y, E)$. For each element $x_i \in A \cup B \cup C$, we add an item coupled-task $x_i$ into $X$ with $\alpha(x_i) = 1$. For each triplet $t_i \in T$, we add a box coupled-task $t_i$ to $Y$ with $\alpha(t_i) = 9$, and an item coupled-task $t'_i$ with $\alpha(t'_i) = 2 + \epsilon$. For each $t_i \in T$ and each $x_i \in t_i$, we add the compatibility arc $(x_i, t_i)$ to $E$. We also add the compatibility arc $(t'_i, t_i)$ to $E$. So, the set of $X$-tasks (resp. $Y$-tasks) are constituted by item coupled-task $x_i$ and $t'_i$ (resp. box coupled-task).

Clearly we have $m$ box coupled-tasks (each with an idle time of 9 units) of degree 4 in $G_c$, $m$ item coupled-tasks with stretch factor $2 + \epsilon$ of degree 1 in $G_c$, and $3n$ item coupled-tasks with stretch factor of degree 2 in $G_c$. Moreover $G_c$ is a bipartite graph.

The reduction is constructed in polynomial time.

It exists a schedule of length $63n - 3k(1 - \epsilon)$ if it exists a matching of size $k$ for $\text{Max-3DM-2}$ instance.

Hereafter, we propose some negative results concerning the existence of subexponential-time algorithms under the following complexity-theoretic hypothesis that is known as the exponential-time hypothesis (see (Woeginger 2001) for a survey on exact algorithms for $\text{NP}$-hard problems) for stretched coupled-tasks, and other ones previously studied.

Recall first the exponential-time hypothesis ((Impagliazzo & Paturi 2001), and (Impagliazzo et al. 2001)): there exists a constant $c > 1$ such that there exists no algorithm for $3$-satisfiability that uses only $O(c^l)$ time where $l$ denotes the number of variables.

A quasi split graph is a connected graph $G = (G_X, G_Y, E)$, with $G_X$ a connected non-oriented graph (not complete) and $G_Y$ a independent set. The other arcs are oriented from $X$ to $Y$ only.
Corollary 1. Assuming the Exponential-Time Hypothesis, there exists no algorithm with a worst-case running time that is subexponential in $n$ (the number of vertices), i.e.:

1. For the 1|ai = bi = p, Li = 2p, Gc|Cmax problem in $O(2^{o(n)})$ time
2. For 1|ai = a, bi = b, Li = a + b, Gc|Cmax in $O(2^{o(n)})$ time
3. 1|α, Gc = 1 - bipartite|Cmax in $O(2^{O(n)})$-time algorithm.

Proof. 1. For 1|ai = bi = p, Li = 2p, Gc|Cmax: In (van Rooij et al. 2013), the authors proved that for PARTITION INTO TRIANGLES on graphs of maximum degree four, there is no algorithm with a worst-case running time $O(2^{o(n)})$ that is subexponential in $n$.

Therefore, we transform a PARTITION INTO TRIANGLES instance with $n$ vertices and $m$ edges into an equivalent instance $G_c$ for bounded degree at most four. Since the transformation is linear (see (Simonin et al. 2012)) the result holds.

2. For the problem 1|ai = a, bi = b, Li = a + b, Gc|Cmax: In (Lokshatnov, Marx & Saurabh 2011) the authors proved that for HAMILTONIAN PATH there is no $O(2^{o(n)})$-time algorithm. As the same way as previously the transformation is linear (see (Simonin et al. 2012)).

3. 1|α, Gc = 1 - bipartite|Cmax: In (Chen, Jansen & Zhang 2014), the authors proved that for MAX 3DM, there is no $O(2^{o(n)})$-time algorithm, therefore this result is transposed to the scheduling problem using the first part of the proof of Theorem 1.

Theorem 2. Scheduling stretched coupled task in presence of a quasi split graph is a $\mathcal{NP}$-complete problem even if $#(V(G_X)) = 1$ and $#(V(G_Y)) = 1$

Proof. The proof is based on a reduction from a variant of the well-know $\mathcal{NP}$-complete PARTITION INTO TRIANGLES. This problem consists to ask if the vertices of a graph $G = (V, E)$, with $|V| = 3q, q \in \mathbb{N}$, can be partitioned into $q$ disjoint sets $T_1, T_2, \ldots, T_q$, each containing exactly three vertices, such that for each $T_i = \{u_i, v_i, w_i\}, 1 \leq i \leq q$, all three of the edges $\{u_i, v_i\}, \{u_i, w_i\}, \{w_i, v_i\}$ belong to $E$.

The problem PARTITION INTO TRIANGLES remains $\mathcal{NP}$-complete even if the graph $G$ can be partitioned into three sets with the same size, $A$, $B$ et $C$ such that each set is an independent set (Morandini, M. 2004). The polynomial-time transformation is based on this variant. Let $G = (A \cup B \cup C, E)$ be an instance of the variant of PARTITION INTO TRIANGLES. We consider the split-graph $G' = (A \cup B, C, E')$ obtained as follows:

For all $v \in A$ (resp. $B$), we create a vertex $A_v$ (resp. $B_v$) with processing time $(1,1,1)$. Moreover, $\forall v \in C$ we create a task $C_v$ with processing time $(4,4,4)$. The edges between $A$ and $B$ remained the same as the $G'$ whereas the edge between $A \cup B$ and $C$ are oriented. Finally in order to have a connected graph, we add two new vertices (resp. one) $z_0$ and $z_1$ (resp. $z_2$ with processing time equal to $(1,1,1)$ (resp. $(4,4,4)$). We add edges between $z_0$ to $A_v$ (resp. $z_1$ to $B_v$). Lastly, we add the three edges $(z_0, z_2), (z_1, z_2)$ and $(z_0, z_1)$.

Notice that the graph $A_v \cup B_v$ form a bipartite graph. The problem is clearly in $\mathcal{NP}$. It exists a positive solution for the variant of PARTITION INTO TRIANGLES iff a valid schedule of length $12 \times ((|C| + 1)$ exists. It is sufficient to execute the two tasks $A_v$ and $B_v$ in four units of time into a task $C_v$.

Theorem 3. The problem is $5/4$-approximable on quasi split-graph where $#(V(G_Y)) = 1$.

Proof. W.l.o.g., we suppose that the processing time of $X$-tasks (resp. $Y$-tasks) is $(1,1,1)$ (resp. $\alpha(y_i)$). Indeed, if $\alpha(x) > 1$, we put $\alpha(y_i) = \left\lfloor \frac{\alpha(y_i)}{\alpha(x)} \right\rfloor$ and $\alpha(x) = 1$.

Algorithm: we transform the problem into an oriented maximum flow-problem between $G_X$ and $G_Y$ with two sources $s$ and $t$, with $\omega(s, x) = \omega(x, y) = 1$ and $\omega(y, t) = \left\lfloor \frac{\alpha(y_i)}{\omega(y, t)} \right\rfloor$, $\forall y_i \in Y, \forall x \in XG_Y$ where $\omega(i, j)$ is the capacity of an arc $(i, j)$ . After the computation of a maximum flow $F$ of value $f$, for the uncovered remaining $X$-tasks a maximum
$M$-matching ($|M| = m$) is applied. The schedule consists in processing first, the $Y$-tasks with $X$-tasks inside. The $M$-tasks are executed after. Lastly, we schedule $s$ isolated-tasks.

The length of schedule given by the algorithm is $C_{\text{max}} \leq \sum_{y \in Y} 3\alpha(y) + 4m + 3s$ with $2m + s + f = n = |X|$ and $\sum_{y \in Y} 3\alpha(y) \geq 9f$. In similar way, the optimal length is $C_{\text{max}}^* \geq \sum_{y \in Y} \alpha(y) + 4m^* + 3s^*$. We suppose that in $Y$-tasks where are $p^*$-edges processed and $r^*$ isolated-tasks, then we obtain $2(p^* + m^*) + r^* + s^* = n$, $p^* + r^* \leq f$, and $\sum_{y \in Y} \alpha(y) \geq 12p^* + 9r^*$. In the worst-case, the $p^*$-edges are split into two tasks (so $p^*$ news tasks are added to $s^*$), and also the matched-edges are split (for each $m^*$ edges one task is executed into the $Y$-task, instead of one of $r^*$-tasks). Therefore, $2m^*$ tasks are added to the $s$-value. In the worst case, we have $m^* = r^*$, $s = s^* + p^* + 2r^*$ and $m = 0$. In such case, $C_{\text{max}}^* \leq 12p^* + 9r^* + 3s^* + 3p^* + 6r^*$ and $C_{\text{max}}^* = 12p^* + 9r^* + 4r^* + 3s^*$. Thus $\rho \leq \frac{15p^* + 15r^* + 3s^*}{12p^* + 13r^* + 3s^*} \leq \max(5/4, 15/13, 1) = 5/4$.

**Tightness:** it exists an example for the $C_{\text{max}} = 36$, and for the heuristic $C_{\text{max}} = 45$. Consider the graph: three triangles $(x_1, x_2, y_1)$, $(x_3, x_4, y_2)$, and $(x_5, x_6, y_1)$. We add the edges $(x_2, y_3, (x_3, y_1)$ and $(x_5, y_2)$. The optimal solution consists in executing the $X$-tasks into the $Y$-tasks; whereas the heuristic leads the solution in which three $X$-tasks are processed after the $Y$-tasks.

**References**

Blazewicz, J., Ecker, K., Kis, T., Potts, C., Tanas, M. & Whitehead, J. (2009), Scheduling of coupled tasks with unit processing times, Technical report, Poznan University of Technology.


