Note

Approximation of the $k$-batch consolidation problem

Sung-Pil Hong$^a$, Myoung-Ju Park$^{a,*}$, Soo Y. Chang$^b$

$^a$ Department of Industrial Engineering, Seoul National University, San 56-1 Shilim-Dong, Kwanahk Gu, Seoul, 151-742, Republic of Korea
$^b$ Department of Industrial and Management Engineering, Pohang University of Science and Technology, San 31 Hyoja-Dong, Pohang, Kyungbuk, 790-784, Republic of Korea

A R T I C L E   I N F O

Article history:
Received 23 June 2008
Accepted 5 November 2008
Communicated by D.-Z. Du

Keywords:
k-batch consolidation problem
Inapproximability
Approximation algorithm
$k$-set cover problem

A B S T R A C T

We consider a problem of minimizing the number of batches of a fixed capacity processing the orders of various sizes on a finite set of items. This batch consolidation problem is motivated by the production system typical in raw material industries in which multiple items can be processed in the same batch if they share sufficiently close production parameters. If the number of items processed in a batch is restricted to being up to some fixed integer $k$, the problem is referred to as the $k$-batch consolidation problem. We will show that the $k$-batch consolidation problem admits an approximation whose factor is twice that of the $k$-set cover problem. In particular, this implies an upper bound on the approximation factor, $2H_k - 1$, where $H_k = 1 + 1/2 + \cdots + 1/k$.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Consider a production system where the orders $r(v) \in \mathbb{Q}_+$ on a finite set of items $v \in V$ are processed in batches. Each batch has a fixed capacity 1: the total order of items processed in a single batch cannot exceed 1. We are given a set, $E \subseteq V \times V$, of pairs of compatible items $(u, v)$. Any set of items $S \subseteq V$ can be processed in the same batch if and only if they are compatible pairwise; in other words, $S$ induces a clique on the compatibility graph $G = (V, E)$. Then naturally we can consider a problem of finding a minimum number of batches that can process the complete set of orders $\{r(v) : v \in V\}$. This problem will be referred to as the batch consolidation problem or generalized batch consolidation problem for emphasis. If there is an additional constraint that each batch cannot process more than $k$ items for some constant $k \in \mathbb{Z}_+$, we call the problem the $k$-batch consolidation problem, which models the situation where proliferation of items in a single batch is prohibitive for a logistic reason.

Consider an integral version of the batch consolidation problem: given integer-valued orders $r(v) \in \mathbb{Z}_+$ and batch capacity $\lambda \in \mathbb{Z}_+$, the orders of items processed in the batches are required to be integer-valued. But, in [1], it has been observed that, when $k = 2$, given an optimal solution when allowed to process non-integral orders, one can construct the solution processing integral orders without increasing the number of utilized batches. Therefore it is an optimal solution of the integral version of the problem. It is not hard to show that such an observation extends to a general $k$. Thus, our definition of the batch consolidation problem using a unit batch capacity is general enough to cover the integral version.

The batch consolidation problem, first proposed by Lee et al. [7], was motivated by a production system typical for raw material industries such as steel, chemical and semiconductor ones. The process of a particular batch is characterized by a finite set of production parameters. Hence multiple items can be processed in the same batch if their parameters are sufficiently close. Naturally, the production efficiency depends on how well the batches are consolidated so that the number of utilized batches is minimized.

* Corresponding author. Tel.: +82 2 884 1164.
E-mail addresses: pmj0684@snu.ac.kr (M.-J. Park), syc@postech.ac.kr (S.Y. Chang).
URL: http://polytope.snu.ac.kr/sphong.html (S.-P. Hong).

0304-3975/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2008.11.007
It is not hard to see that the batch consolidation problem includes the clique partition problem [7,1], which implies that it does not admit an approximation within a factor of $|V|^\epsilon$ for some $\epsilon > 0$ [9]. But, as we will see, the $k$-batch consolidation problem is approximable within $2H_k - 1$ times the optimum, where $H_k = 1 + \frac{1}{2} + \cdots + \frac{1}{k}$. The idea is to decompose the orders of items so that a minimum cardinality set cover problem whose elements of the ground set correspond to the decomposed orders provides a well-consolidated set of batches. Note that this algorithm provides a 2-approximation when $k = 2$. Chang et al. [1] develop a $\frac{3}{2}$-approximation algorithm, for $k = 2$, based on a more elaborate scheme for the decomposition of orders of items. However, as also will be discussed later, once $k$ becomes $> 3$ such a scheme does not help in improving the approximation factor to strictly better than the one provided by the algorithm proposed in this paper.

The batch consolidation problem is related to bin packing with conflicts, or BPC [6,4]. Although a bin packing problem is fundamentally different from the batch consolidation problem as an item cannot be split over bins, BPC bears some similarity with the batch consolidation problem in that it specifies the pairs of items that cannot be packed in the same bin. [1] discusses some relations between the batch consolidation problem and BPC.

Another related model is packing splittable items with cardinality constraints, or PSIC [5]. PSIC is a generalization of bin packing: the items can be split over bins but a bin cannot contain more than $k$ items. Notice, then, that PSIC [5] is the special case of the $k$-batch consolidation problem in which the compatibility graph $G$ is complete. In this special case, the problem admits a polynomial time approximation scheme while, in general, the problem is max-SNP-hard and not approximable within 1.0021 times the optimum as discussed later.

This paper is organized as follows. Section 2 discusses a simple but useful property of an optimal solution helpful in the analysis of the approximation algorithm. In Section 3, we establish an inapproximability of the $k$-batch consolidation problem. Section 4 is devoted to the discussion of an approximation algorithm.

2. Preliminaries

Given a solution of the problem, consider the hypergraph $\mathcal{H} = (V, B)$ whose vertices and edges, respectively, correspond to the items $V$ and the collection of batches $B \in B$ processing items with their nonzero orders. (See Fig. 1.)

**Proposition 2.1.** Any solution of the $k$-batch consolidation problem can be modified efficiently without increasing the number of utilized batches so that its hypergraph $\mathcal{H}$ is acyclic: there is no sequence $(v_1, B_1, v_2, B_2, \ldots, v_l, B_l, v_1)$ with $l \geq 2$ such that $B_i$ are all distinct, $v_i$ are all distinct, and $v_i, v_{i+1} \in B_i$ for $i = 1, 2, \ldots, l - 1$, and $v_1, v_l \in B_l$. (We refer to such sequence as a circuit of a hypergraph.)

**Proof.** Suppose $\mathcal{H}$ has a circuit $C := (v_1, B_1, v_2, B_2, \ldots, v_l, B_l, v_1)$. For each $v_i$, $i = 1, 2, \ldots, l$, from $C$, write, as $\rho_i^-$ and $\rho_i^+$, respectively, the orders of $v_i$ processed by $B_{i-1}$ and $B_i$ batches where $B_0$ denotes $B_l$. Reverse the direction and/or redesignate the initial vertex of circuit so that $\min_{v_i \in C} \{\rho_i^-, \rho_i^+\} = \rho_i^+$. Then the modified solution $\rho_i^+ - \rho_i^-, \rho_i^+, \rho_i^- + \rho_i^+, \forall v_i \in C$ is feasible with the same number of batches. And the circuit is disconnected since $v_1$ is deleted from $B_l$. Repeating this procedure, we can convert any solution to having an acyclic hypergraph. \(\square\)

**Proposition 2.2.** Any problem $(G, r)$ can be reduced in polynomial time into one, $(G, s)$, with $s(v) < \deg(v) + 1$. $\forall v \in V$.

**Proof.** From Proposition 2.1, there is an optimal solution which has at most $\deg(v)$ batches processing the order of $v \in V$ with other items. In other words, when $r(v) \geq \deg(v) + 1$, we can first construct $\lceil r(v) - \deg(v) \rceil$ batches processing exactly 1 out of the order $r(v)$ without compromising the optimality. Then the orders are reduced to $s(v) = r(v) - \lceil r(v) - \deg(v) \rceil < \deg(v) + 1$. \(\square\)

3. Inapproximability

We can derive an easy inapproximability of the $k$-batch consolidation problem from the inapproximability of the 2-batch consolidation problem by Chang et al. [1]. In the reduction from the vertex cover problem with bounded degree to an instance of the 2-batch consolidation problem, they construct the compatibility graph $G$ as bipartite. A $k$-batch consolidation problem on a bipartite $G$ is simply a 2-batch consolidation problem and hence the reduction is also valid for the $k$-batch consolidation problem.

**Theorem 3.1.** The $k$-batch consolidation problem cannot be approximated within 1.0021 times the optimum for all $k \geq 2$ unless $P = NP$. 

---

**Fig. 1.** The hypergraph determined by the batches $B_1, B_2$ and $B_3$. 

![Diagram](image-url)
4. Approximation

4.1. Set-cover-based algorithm

For a problem \((G, r)\) defined by \(G = (V, E)\) and \(r \in \mathbb{Q}^+\), the approximation algorithm is conveniently described by defining an auxiliary problem \((U, \delta)\) where \(U\) and \(\delta\) are constructed as follows. For each \(v \in V\), compute \(n_v := |kr(v)|\) and accordingly construct a decomposition of \(r(v)\): a set of \(n_v\) elements, \(D_v = \{v_0, v_1, v_2, \ldots, v_{n_v-1}\}\) and their orders \(r'(v_0) = r(v) - \frac{n_v - 1}{\lambda}, r'(v_1) = \cdots = r'(v_{n_v-1}) = \frac{1}{\lambda}\). Let \(U = \bigcup_{v \in V} D_v\) and \(\delta = \{S \subseteq U | \sum_{u \in S} r'(u) \leq 1, \delta \leq k, S\) is a clique\). The following \((U, \delta)\) is an auxiliary problem of instance \(G = (V, E)\) in Fig. 1 for \(k = 2\).

Algorithm 4.1.

Step 1 Construct the auxiliary problem \((U, \delta)\) of \((G, r)\).

Step 2 Compute a minimum set cover \(C \subseteq \delta \cup U\).

Step 3 For each subset \(C\) of \(E\), construct a batch processing the assigned orders, \(r'(u), \forall u \in C\). Return the batches as a solution.

From the construction of \((U, \delta)\), the batches from Step 3 can cover the complete set of orders. Also notice that we can adjust the processing orders without increasing the number of batches so that each order \(r'(u)\) is exactly covered. As the orders \(r'\) of the auxiliary problem are a decomposition of the original orders \(r\), the batches from Step 3 are clearly a feasible solution of the original problem.

Let us define some more notation. Denote by \(OPT(G, r)\) and \(z(G, r)\), respectively, the numbers of batches of an optimal solution and a solution returned by Algorithm 4.1. And \(c(X)\) is the cardinality of a minimum set cover from \(\delta\) of \(X \subseteq U\). Then, for the analysis of Algorithm 4.1, the following lemmas are useful.

Lemma 4.2. For any partition \((X; Y)\) of \(U\), \(z(G, r) = c(U) \leq c(X) + c(Y)\).

Proof. Let \(C_1\) and \(C_2\), respectively, be the minimum set covers of \(X\) and \(Y\). Then, \(C_1 \cup C_2\) is a set cover of \(U\). Therefore, \(z(G, r) = c(U) \leq |C_1| + |C_2| = c(X) + c(Y)\).

Lemma 4.3. If \(r \leq s\), \(OPT(G, r) \leq OPT(G, s)\) and \(z(G, r) \leq z(G, s)\).

Proof. The first half of the statement is trivial.

For the second half, define \(r'(w) = s(w) - \delta\), with \(0 < \delta \leq \frac{1}{\lambda}\) for any fixed \(w \in V\) and \(r' = s, \forall v \in V \setminus \{w\}\). Let \((U_r, \delta_r)\) and \((U_s, \delta_s)\), respectively, be the auxiliary problems for \((G, r)\) and \((G, s)\). Also let \(D_w = \{w_0, w_1, \ldots, w_i\}\) be a set for \(w\) in the decomposition of \(s(w)\) from \((G, s)\).

If \(s'(w_0) > \delta\) in the decomposition of \(s(w)\), then \(U_r = U_s\) and hence \(\delta_s \supseteq \delta_r\). Therefore a set cover of \((U_s, \delta_s)\) is also a set cover of \((U_r, \delta_r)\) and we have \(z(G, r) \leq z(G, s)\).

If, on the other hand, \(s'(w_0) \leq \delta\), then we get \(D_w = \{w_0, w_1, \ldots, w_{i-1}\}\) for \(w\) in the decomposition of \(r'(w)\) from \((G, r)\).

Let \(C_r\) be any set cover of \(U_r\). Delete \(w_0\) from the subset of \(C_r\). And replace \(w_i\) with \(w_0\). Then the resulting collection is a set cover of \(U_s\), whose cardinality is no greater than \(|C_r|\). (See Fig. 2.) Therefore we have \(z(G, r) \leq z(G, s)\).

Theorem 4.4. \(z(G, r) \leq 2OPT(G, r)\).

Proof. By induction on \(OPT(G, r)\). Suppose \(OPT(G, r) = 1\). Then \(G\) is a complete graph with \(|V| \leq k\) and \(\sum_{v \in V} r(v) \leq 1\). This implies that both \(S_1 = \{u \in U| r'(u) < \frac{1}{\lambda}\}\) and \(S_2 = \{u \in U| r'(u) = \frac{1}{\lambda}\}\) are elements of \(\delta\). But, \(S_1, S_2\) is a set cover of \(U\) and we have \(z(G, r) \leq 2 = 2OPT(G, r)\).
Proposition 4.4

Lemma 9.6


can, however, rely on an approximationalgorithm of the $k$-set cover problem. For instance, the approximationalgorithm of [2] guarantees the approximation factor of $(H_k - \frac{1}{2})$, where $H_k = 1 + \frac{1}{2} + \cdots + \frac{1}{k}$. Thus, employing the approximate solution
instead of an exact one, Algorithm 4.1 is an \((2H_k - 1)\)-approximation. For \(k \geq 4\), we can use the \((H_k - \frac{196}{195})\)-approximation algorithm of [8] instead to get a slightly improved \((2H_k - \frac{196}{195})\)-approximation for the \(k\)-batch consolidation problem.

4.2. An alternative decomposition scheme

When \(k = 2\), Algorithm 4.1 employing an \((H_k - \frac{1}{2})\)-approximation of the \(k\)-set cover problem provides a 2-approximation as the corresponding 2-set cover problem is nothing but the polynomially solvable minimum edge-cover problem. But, the specialized 2-batch problem algorithm of [1] guarantees the approximation factor, \(\frac{1}{2}\). The algorithm is based on the same idea of solving the edge-cover problem on the auxiliary problem obtained by decomposition of the orders of vertices. But, it uses a slightly different decomposition: each vertex \(v\) of order \(r(v)\) is decomposed into \(2 \times \lfloor \frac{1}{2} r(v) \rfloor\) vertices all assigned the order \(\frac{1}{2}\), and one vertex of the order \(r(v) - \lfloor \frac{1}{2} r(v) \rfloor\) in the auxiliary problem. The remaining steps are exactly the same as Algorithm 4.1 for \(k = 2\). Thus there is one (and at most one) auxiliary vertex per original one, whose order can be greater than \(\frac{1}{2}\). As shown in [1], when the number of items processed in a single batch is restricted to a number as small as \(k = 2\), such vertices are crucial in attaining the approximation guarantee of \(\frac{1}{2}\).

Interestingly enough, when \(k \geq 3\), however, such a decomposition scheme does not help in improving the approximation guarantee to strictly better than 2. To see this, consider the complete graph \(G = (V, E)\) with \(|V| = 2l - 1\) and \(r(v) = \frac{1}{2} + \varepsilon\), \(\forall v \in V\). Then, for each \(v \in V\), we get \(r'(v_0) = \frac{1}{2} + \varepsilon\) and therefore \(v_0\) participates only in a singleton set in the auxiliary problem. Thus \(z(G, r) = 2l - 1\) while the optimal value of the \(k\)-batch problem is \(OPT(G, r) = \lceil (2l - 1)(\frac{1}{2} + \varepsilon) \rceil = l\) for all \(k \geq 3\). Thus the approximation factor is \(2 - \frac{1}{2}\).

5. Further research

There is currently a significant gap between the upper bound \((2H_k - 1)\) and the lower bound \(1.0021\) on the approximability of the \(k\)-batch consolidation problem. It is an interesting open problem whether the lower bound can be tightened to \(\log k\), asymptotically the same as the upper bound, or vice versa.

Acknowledgement

The second author’s research was partially supported by the Second Stage of Brain Korea 21 Project in 2008.

References