Approximability of a batch consolidation problem

Junho Chang, Soo Y. Chang
Department of Industrial Management Eng., Postech, Korea
Sung-Pil Hong, Yun-Hong Min, and, Myoung-Ju Park
Department of Industrial Eng., Seoul National University, Korea

November, 14, 2008
Discrete Math. Seminar, KAIST

Consider a production system where orders are processed in batches.

- Orders $r(v) \in \mathbb{Z}_+$ on a finite set of items $v \in V$.
- Each batch has a fixed capacity $\lambda \in \mathbb{Z}_+$.
- Pairs of items $(u, v) \in E \subseteq V \times V$ that can be processed in the same batch.
- Each batch can process up to 2 items, 2-BCP.

To find a minimum number of batches covering the orders.
Modeled by

General case is harder than CLIQUE PARTITION; inapproximable within $|V|^{\epsilon}$ for some $\epsilon > 0$.

$k$-BCP admits approximation of factor, $2H_k - 1$, where $H_k = 1 + \frac{1}{2} + \cdots + \frac{1}{k}$.

Bin packing with conflicts bears some similarity.
Problem instance $\leftrightarrow$ a pair, $(G, r)$.

Feasible solution will be denoted by $S(G, r)$

*Pure, mixed, saturated, and nonsaturated* batches.

*Solution graph* of $S(G, r)$ is graph defined on $V$. Two nodes $u, v$ are connected as many times as numbers of mixed batches processing $u$ and $v$. 
Number of batches used by feasible solution $S(G, r)$ is denoted by $z_S(G, r)$

Feasible solution $S^*(G, r)$ giving smallest possible $z_S(G, r)$ is called *optimal* solution and $z_{S^*}(G, r)$ is denoted by $\text{OPT}$.

If algorithm $A$ guarantees $S(G, r)$ with $z_S(G, r) \leq \rho \text{OPT}$, we call $A$ a $\rho$-approximation algorithm, and $\rho$ *approximation factor* or *approximation ratio*. 
Proposition

Given any feasible solution, we can find a solution in polynomial time satisfying the followings.

1. For each edge, there is at most one mixed batch.
2. For each vertex, there is at most one nonsaturated pure batch.

Proof

First part: Suppose \( uv \in E \) has two mixed batches.

Second part: trivial. \( \square \)
‘Degree+1 Proposition’

**Proposition**

*Any problem* \((G, r)\) *can be reduced in polynomial time into one* \((G, r')\) *with* \(r'(v) < (\text{deg}(v) + 1)\lambda, \forall v \in V.\)

**Proof**

- From previous Proposition, there is optimal solution using at most \(\text{deg}(v)\) mixed batches for \(v \in V.\)
- When \(r(v) \geq (\text{deg}(v) + 1)\lambda\), we first construct \(\left\lfloor \frac{r(v) - \lambda \text{deg}(v)}{\lambda} \right\rfloor\) saturated pure batches out of \(r(v)\) without compromising optimality.
- Then orders are reduced to \(r'(v) = r(v) - \lambda \left\lfloor \frac{r(v) - \lambda \text{deg}(v)}{\lambda} \right\rfloor < (\text{deg}(v) + 1)\lambda.\) \(\square\)
‘No-Circuit Proposition’

**Proposition**

*Any solution can be modified in polynomial time without increasing number of batches so that its solution graph is forest.*

**Proof**

- Suppose solution graph has circuit $C$. Then give an orientation on $C$.

- For $v \in V(C)$, write as $\rho_+(v)$ and $\rho_-(v)$, respectively, orders processed by batches corresponding to outgoing and incoming arcs to $v$.

- Then solution $\rho_+(v) + \epsilon$, $\rho_-(v) - \epsilon \ \forall v \in V(C)$ is also feasible if $\rho_+(v) + \epsilon \geq 0$, $\rho_-(v) - \epsilon \geq 0 \ \forall v \in V(C)$.

- Set $\epsilon = \min\{\rho_-(v)|v \in V(C)\}$ Then at least one arc on $C$ will be deleted. Repeating procedure, we can make solution graph acyclic. □
Proposition

Any solution can be modified in polynomial time without increasing its number of batches so that processed orders are all integers.

Proof Omitted.
From vertex cover with bounded degree, VCBD

Given instance $I$ of VCBD with $G = (V, E)$ and $\delta$ construct instance of BCP, $I' = (G' = (V', E'), r)$ with capacity $\lambda$ as follows:

- $\lambda = \delta + 1$.

- $G'$ is a subdivision of $G$:
Objective-Value Proposition

I has a vertex cover $C$ with $|C| \leq k$ if and only if $I'$ has feasible set of batches whose cardinality $\leq 2|E| + k$. In other words,

$$\text{OPT}_{\text{VCBD}} + 2|E| = \text{OPT}_{\text{BCP}}.$$
OPT_{VCBD} + 2|E| \geq OPT_{BCP}
\[ \text{OPT}_{\text{VCBD}} + 2 |E| \leq \text{OPT}_{\text{BCP}} \]
Inapproximability ratio 1.0021 for BCP

Suppose BCP approximable with factor \((1 + \epsilon)\) and \(S(G', r)\) is such solution of obj value \(z_S\), number of used batches. Let \(C\) be corresp. VC of VCBD. Then

\[
\epsilon \cdot OPT(I') \geq z_S - OPT(I') \\
= |C| + 2|E| - OPT(I') \text{ (from proof of Proposition)} \quad (1) \\
\geq |C| - OPT(I) \text{ (from Proposition)}. 
\]

As \(|E| \leq \delta \cdot OPT(I)|\), Proposition implies \(OPT(I') \leq (2\delta + 1)OPT(I)\) which, combined with (1), in turn, implies

\[
|C| - OPT(I) \leq \epsilon(2\delta + 1)OPT(I). 
\]

Thus, \((1 + \epsilon)\)-approx. of BCP implies \((1 + \epsilon(2\delta + 1))\)-approx. of VCBD. However, for \(\delta = 4\), VCBD is inapprox. within \(\frac{53}{52}\) of optimum unless \(P = NP\). Therefore, for any

\[
\epsilon < \frac{1}{52} \cdot \frac{1}{9} (> 0.0021)
\]

\((1 + \epsilon)\)-approximation is impossible for BCP.
Auxiliary problem \((G_r, r')\) for given \((G, r)\)

- \(V(G_r)\): For \(v \in V\), compute \(k = \lceil r(v)/\lambda \rceil - 1\) and construct \(2k + 1\) vertices, \(V'_v = \{v_0, v_1, v_2, \ldots, v_{2k}\}\). Then \(V(G_r) = \bigcup_{v \in V} V'_v\).

- \(r'\): \(r'(v_0) = r(v) - k\lambda, r'(v_1) = r'(v_2) = \cdots = r'(v_{2k}) = \lambda/2\).

- \(E(G_r)\): For \(v \in V\), construct \(E'_v = \{xy | r'(x) + r'(y) \leq \lambda, x \neq y, \text{ and } x, y \in V'_v\}\). Also for \(uv \in E\), consider edges \(E'_{uv} = \{xy | x \in V'_u, y \in V'_v\}\) and \(r'(x) + r'(y) \leq \lambda\). Then \(E(G_r) = \left(\bigcup_{v \in V} E'_v\right) \cup \left(\bigcup_{uv \in E} E'_{uv}\right)\).
Decomposition of orders

Our decomposition

\[ \begin{array}{cccc}
V_{2k} & \cdots & V_2 & V_1 & V_0 \\
\end{array} \]

\[ \rightarrow 3/2\text{-approximation} \]

All-\(\lambda/2\) decomposition

\[ \begin{array}{cccc}
\hline
\hline
\end{array} \]

\[ \rightarrow 2\text{-approximation} \]
Step 0  Remove pure batches, if any, from each vertex to satisfy 
\[ r(v) < (\deg(v) + 1)\lambda, \forall v \in V. \]

Step 1  Construct auxiliary problem \((G_r, r')\) of \((G, r)\).

Step 2  Compute maximum matching \(M\) on \(G_r\).

Step 3  For each \(M\)-exposed vertex \(y\) of \(G_r\), construct a pure  
batch processing \(r'(y)\) of the order of corresponding item.  
For each edge \(xy \in M\), construct a mixed batch processing  
the orders \(r'(x)\) and \(r'(y)\) of corresp. original items.

Step 4  Return \((|V(G_r)| - 2|M|)\) pure batches and \(|M|\) mixed  
batches resulting from Step 3.
Theorem

$$z(G, r) \leq \frac{3}{2} \text{OPT}(G, r).$$
Preliminaries

For partition \((S; T)\) of \(V(G_r)\), denote by \(r'_S\) and \(r'_T\), resp., \(r'\) restricted to \(S\) and \(T\): 
\[
r'_S(x) = r'(x)\chi^S(x), \quad r'_T(x) = r'(x)\chi^T(x), \quad \forall x \in V(G_r).
\]

‘Partition Lemma’

Lemma

For any partition \((S; T)\) of \(V(G_r)\), 
\[
z(G, r) = z(G_r, r') \leq z(G_r, r'_S) + z(G_r, r'_T).
\]

Proof If we apply Algorithm to \((G_r, r'_S)\) and \((G_r, r'_T)\), respectively, then the two returned solutions correspond to maximum matchings \(M_1\) of \(G_r[S]\) and \(M_2\) of \(G_r[T]\). Then, \(M_1 \cup M_2\) is a matching of \(G_r\). 
\[
z(G, r) = |V(G_r)| - |M| \quad \text{where} \quad M \text{ maximum matching}
\leq |V(G_r)| - (|M_1 \cup M_2|)
= (|S| - |M_1|) + (|T| - |M_2|)
= z(G_r, r'_S) + z(G_r, r'_T). \square
For partition \((S; T)\) of \(V(G_r)\), define \(\sigma\) and \(\tau \in \mathbb{Q}^V_+\) as follows:
\[
\sigma(v) = \sum_{x \in S \cap V'_v} r'_S(x), \quad \tau(v) = \sum_{x \in T \cap V'_v} r'_T(x), \quad \forall v \in V.
\]
(Then \(\sigma + \tau = r\).)
Also denote \(V_S = \{v \in V | S \cap V'_v \neq \emptyset\}\).

\[
\begin{array}{cccc}
\sigma & 1 & 4 & 0 \\
\tau & 0 & 1 & 2 \\
r & 1 & 5 & 2 \\
\end{array}
\]

\(V_S = \{u, v\} \quad V_T = \{v, w\}\)
‘Intersection Lemma’

Lemma

Suppose $V_S \cap V_T = \{v\}$ for some $v \in V$. If $\sigma(v) = k\lambda$ for some $k \in \mathbb{Z}_+$ or $v_0 \in S$, then $z(G_r, r'_T) \leq z(G, \tau)$. ($k = 0$ implies $V_S \cap V_T = \emptyset$.)

Proof: If $\sigma(v) = k\lambda$ for some $k \in \mathbb{Z}_+$ then $(G_r, r'_T)$ is auxiliary problem of $(G, \tau)$ and therefore $z(G_r, r'_T) = z(G, \tau)$.

Suppose $S \ni v_0$. If $T \cap V'_v = \{v_1, \ldots, v_q\}$, then $r'_T(v_1) = \cdots = r'_T(v_q) = \lambda/2$ and $\tau(v) = q \times \frac{\lambda}{2}$.

If $q$ is odd, then, $(G_r, r'_T) = (G_\tau, \tau')$ over nonzero-order vertices and $z(G_r, r'_T) = z(G, \tau)$.
Preliminaries (cont’d)

If \( q \) is even, then, in \((G_\tau, \tau')\), \( v \) is decomposed into \( q - 1 \) vertices, one with order \( \lambda \) and remaining \( q - 2 \) with order \( \lambda/2 \).

Thus, only difference between \((G_r, r'_\tau)\) and \((G_\tau, \tau')\) is the latter has a vertex with order \( \lambda \) instead of two vertices, say, \( v_{q-1} \) and \( v_q \) of the former both with order \( \frac{\lambda}{2} \).

The vertex with order \( \lambda \) in \((G_\tau, \tau')\) will be isolated, but the rest, in their inducing subgraph, inherit adjacency from \((G_r, r'_\tau)\). Thus, if \( M(G_\tau) \) is a maximum matching of \( G_\tau \), \( M = M(G_\tau) \cup \{v_{q-1}v_q\} \) is also a matching of \((G_r, r'_\tau)\). Therefore we get

\[
z(G_r, r'_\tau) \leq |T| - |M| = (|V(G_\tau)| + 1) - (|M(G_\tau)| + 1) = z(G, \tau). \quad \square
\]
‘Incremental Lemma’

Lemma

For fixed \( w \in V \), define \( s = r + \Delta \chi^w \) for constant \( 0 < \Delta \leq \lambda \). Then, \( \text{OPT}(G, r) \leq \text{OPT}(G, s) \leq \text{OPT}(G, r) + 1 \) and \( z(G, s) \leq z(G, r) + 1 \).

Proof First half is trivial.

Consider two subcases of \( 0 < r'(w_0) + \Delta \leq 2\lambda \) in auxiliary problem.

Case 1: \( r'(w_0) + \Delta \leq \lambda \). Then, \( V(G_s) = V(G_r) \). Notice no pure batch is constructed from Step 0 of Algorithm applied to \((G, s)\). Also as, due to increase in \( r'(w_0) \), \( G_s \) has fewer edges incident to \( w_0 \) than \( G_r \), their maximum matchings \( M(G_s) \) and \( M(G_r) \) from Step 2 satisfy \( |M(G_s)| \geq |M(G_r)| - 1 \). Therefore we get

\[
z(G, s) = |V(G_s)| - |M(G_s)| \leq |V(G_r)| - (|M(G_r)| - 1) = z(G, r) + 1.
\]

Case 2: \( \lambda < r'(w_0) + \Delta \leq 2\lambda \). Omitted. \( \square \)
Sketch of Proof

By induction on $OPT(G, r)$.
For $OPT(G, r) = 1$, easy to show $z(G, r) = 1$.
Suppose $OPT(G, r) = n$ with opt. solution $S^*(G, r)$. If $S^*(G, r)$ has pure batch or if its solution graph has component of single edge, induction is easy.
Otherwise, solution graph, as forest, has path $P = u - v - w$ of length 2 in which
$\deg(u) = 1$ and $\deg(v) = 2$, type I or
$\deg(u) = 1, \deg(v) \geq 3$ and $\deg(w) = 1$, type II.

---

Junho Chang, Soo Y. Chang  
Department of Industrial Management Eng., Postech, Korea  
Sung-Pil Hong, Yun-Hong Min, and Myoung-Ju Park  
Department of Industrial Eng., Seoul National University, Korea

Approximability of a batch consolidation problem
Let $p$ be vector of orders processed by two mixed batches of path $P$. Notice $p(u) + p(v) + p(w) \leq 2\lambda$. And let $q$ be remaining order (i.e., $q = r - p$). Notice $2 + OPT(G, q) = OPT(G, r)$.

Will find partition $(S; T)$ of $V(G_r)$ satisfying conditions of ‘Intersection’ Lemma and relation ($\ast$):

$$z(G_r, r'_S) + z(G, \tau) \leq 3 + \frac{3}{2}OPT(G, q). \quad (\ast)$$

Then it follows that

$$z(G, r) \leq z(G_r, r'_S) + z(G_r, r'_T) \leq z(G_r, r'_S) + z(G, \tau) \leq 3 + \frac{3}{2}OPT(G, q) = \frac{3}{2}(2 + OPT(G, q)) \leq \frac{3}{2}OPT(G, r).$$

where, first three ineq.’s in their order follow from ‘Partition Lemma’, ‘Intersection Lemma, ($\ast$), and ‘Path Lemma’.
Sketch of Proof: Type I (cont’d)

- Set $\Delta = p(w)$. (Notice $p(w) \leq \lambda$.)
- Case 1: $p(v) \leq \lambda$. Then set $S = \{u_0, v_0\}$.
- Case 2: $\lambda < p(v) (< 2\lambda)$. Then set $S = \{u_0, v_0, v_1, v_2\}$.

Notice in both, conditions of Intersection Lemma are satisfied. Regarding ($\ast$),

$$z(G_r, r'_S) + z(G, \tau) \leq 2 + z(G, q + \Delta \chi^w) \leq 3 + z(G, q) \leq 3 + \frac{3}{2} \text{OPT}(G, q),$$

where, first ineq. follows from $z(G_r, r'_S) \leq 2$, second from Incremental Lemma, and third from induction hypothesis.
Case 1: $p(v) \leq \lambda$. Then set $S = \{u_0, w_0\}$ and $\Delta = p(v)$.

Case 2: $\lambda < p(v) (< 2\lambda)$ and $p(u), p(w) \leq \lambda/2$. Then, set $S = \{u_0, v_1, v_2, w_0\}$ and $\Delta = p(v) - \lambda$.

Arguments are analogous to the ones for type I and are omitted.

Case 3: $\lambda < p(v) (< 2\lambda)$, $p(u) > \lambda/2$, and $p(w) \leq \lambda/2$.

For this last case, we consider three subcases according to whether $v_0$ is connected to $u_0$ and/or $w_0$. 
Sketch of Proof: Type II (cont’d)

In all cases, \( z(G_r, r_S') = 3 \). Also notice conditions of Intersection Lemma are satisfied for all. Regarding (*),

\[
z(G_r, r_S') + z(G, \tau) = 3 + z(G, \tau) \leq 3 + \frac{3}{2} \text{OPT}(G, \tau) \leq 3 + \frac{3}{2} \text{OPT}(G, q),
\]

where, first ineq. follows from induction hypothesis.

Junho Chang, Soo Y. Chang  Department of Industrial Management Eng., Postech, Korea
Sung-Pil Hong, Yun-Hong Min, and Myoung-Ju Park  Department of Industrial Eng., Seoul National University, Korea

Approximability of a batch consolidation problem
Sketch of Proof: Type II (cont’d)

Second ineq. is from $\tau \leq q$ or equiv. $\sigma \geq p$. To see this, first notice that in all cases, $\sigma(u) = p(u) = r'(u_0)$ and $\sigma(w) = p(w) = r'(w_0)$. Therefore if $\sigma(u) + \sigma(v) + \sigma(w) = \sum_{x \in S} r'(x) > 2\lambda$, then combined with $2\lambda \geq p(u) + p(v) + p(w)$ it implies $\sigma(v) \geq p(v)$. The premise $\sum_{x \in S} r'(x) > 2\lambda$ is easily checked in first two subcases. Arguments for third subcase is omitted. $\square$
All-$\frac{\lambda}{2}$ decomposition can not attain factor $\frac{3}{2}$

However, this decomposition yields $(2H_k - 1)$-approximation for $k$-BCP. Interestingly, the decomposition hired for 2-BCP does not essentially improve ratio of all-$\frac{\lambda}{2}$ decomposition when $k \geq 3$. 
Further research

- Online version?
- Still a significant gap: 1.0021 vs. 1.5.
- Better approximation or tighter lower bound for $k$-BCP: 1.0021 vs. $2H_{k-1}$. 

Junho Chang, Soo Y. Chang  
Department of Industrial Management Eng., Postech, Korea  
Sung-Pil Hong, Yun-Hong Min, and Myoung-Ju Park  
Department of Industrial Eng., Seoul National University, Korea