When LP is the Cure for Your Matching Woes: Improved Bounds for Stochastic Matchings

Jian Li, University of Maryland, College Park

Joint work with Nikhil Bansal (IBM), Anupam Gupta (CMU), Julian Mestre (MPI), Viswanath Nagarajan (IBM), Atri Rudra (SUNY-Buffalo)
**Problem Definition**

**Stochastic Matching** [Chen, Immorlica, Karlin, Mahdian, and Rudra. ’09]

- **Given:**
  - A probabilistic graph $G(V,E)$.
  - Existential prob. $p_e$ for each edge $e$.
  - Patience level $t_v$ for each vertex $v$.

- **Probing** $e=(u,v)$: The only way to know the existence of $e$.
  - We can probe $(u,v)$ only if $t_u>0$, $t_v>0$.
  - If $e$ indeed exists, we should add it to our matching.
  - If not, $t_u = t_u - 1$, $t_v = t_v - 1$. 
**Problem Definition**

- **Output:** A strategy to probe the edges
  - Edge-probing: an (adaptive or non-adaptive) ordering of edges.
  - Matching-probing: \( k \) rounds; In each round, probe a matching.

- **Objectives:**
  - Unweighted: Max. \( E[\text{cardinality of the matching}] \).
  - Weighted: Max. \( E[\text{weight of the matching}] \).
MOTIVATIONS

- **Online dating**
**Motivations**

- **Online dating**
  - **Existential prob.** $p_e$ : estimation of the success prob. based on users’ profiles.
  - **Probing edge** $e=(u,v)$ : $u$ and $v$ are sent to a date.
Motivations

- **Online dating**
  - **Existential prob.** $p_e$: estimation of the success prob. based on users’ profiles.
  - **Probing edge** $e=(u,v)$: $u$ and $v$ are sent to a date.
  - **Patience level**: obvious.
Motivations

- Kidney exchange
  - Existential prob. $p_e$ : estimation of the success prob. based on blood type etc.
  - Probing edge $e=(u,v)$ : the crossmatch test (which is more expensive and time-consuming).
OUR RESULTS

Previous results for unweighted version [Chen et al. ’09]:
- Edge-probing: Greedy is a 4-approx.
- Matching-probing: O(log n)-approx.

A simple 8-approx. for weighted stochastic matching.
- For edge-probing model.
- can be improved to 5.75 by a more careful analysis.

An improved 3-approx. for bipartite graphs and 4-approx. for general graphs based on dependent rounding [Gandhi et al. ’06].
- For both edge-probing and matching-probing models.
- This implies the gap between the best matching-probing strategy and the best edge-probing strategy is a small const.
OTHER RESULTS

- Stochastic online matching.
  - A set of items and a set of buyer types. A buyer of type $b$ likes item $a$ with probability $p_{ab}$. (G(items, buyer types): Expected graph)
  - The buyers arrive online (her type is an i.i.d. r.v.).
  - The algorithm shows the buyer (of type $b$) at most $t_b$ items one by one.
  - The buyer buys the first item she likes or leaves without buying.
  - This generalizes the stochastic online matching problem of [Feldman et al. ’09, Bahmani et al. ’10, Saberi et al ’10] where $p_e = \{0,1\}$.

  - We have a 7.92-approximation.
OTHER RESULTS

- Cardinality Constrained Matching in Rounds.
  - In each round, we can probe a matching of size \( \leq C \).
  - An \( O(1) \)-approx.
  - Chen et al. obtained an \( O(\min(k,C)) \)-approx.

- A new proof for greedy.
  - An simple LP-based analysis: 5-approx.
  - The analysis by Chen et al. was based on decision trees.
Other Results

- Stochastic $k$-set packing.
  - Generalizing the stochastic matching problem.
  - $k=4$.

$\begin{align*}
x_e &= 1 \text{ w.p. } p_e \\
&= 0, \text{ o.w.}
\end{align*}$

$\begin{align*}
\begin{bmatrix}
1 \\
1 \\
\vdots \\
t_1 \\
t_2 \\
\vdots
\end{bmatrix}
\end{align*}$

Capacity vector

$\begin{align*}
u: & \quad x_e \\
v: & \quad x_e \\
u: & \quad 1 \\
v: & \quad 1
\end{align*}$

Packing

$n$ items/columns
**Approximation Ratio**

- We compare our solution against the optimal (adaptive) strategy (not the offline optimal solution).

- An example:

  \[ p_e = \frac{1}{n} \]

  \[ E[\text{offline optimal}] = 1 - (1 - \frac{1}{n})^n \approx 1 - \frac{1}{e} \]

  \[ E[\text{any algorithm}] = \frac{1}{n} \]
**A LP Upper Bound**

- **Variable** $y_e$: Prob. that any algorithm probes $e$.

Maximize

$$\sum_{e \in E} w_e \cdot x_e$$

subject to

$$\sum_{e \in \partial(v)} x_e \cdot 1 \quad \forall v \in V$$

At most 1 edge in $\partial(v)$ is matched

$$\sum_{e \in \partial(v)} y_e \cdot t_v \quad \forall v \in V$$

At most $t_v$ edges in $\partial(v)$ are probed

$$x_e = p_e \cdot y_e \quad \forall e \in E$$

$x_e$: Prob. $e$ is matched

$$0 \cdot y_e \cdot 1 \quad \forall e \in E$$
A SIMPLE 8-APPROXIMATION

An edge \((u,v)\) is \textit{safe} if \(t_u > 0, t_v > 0\) and neither \(u\) nor \(v\) is matched

Algorithm:
- Pick a permutation \(\pi\) on edges uniformly at random
- For each edge \(e\) in the ordering \(\pi\), do:
  - If \(e\) is not safe then do not probe it.
  - If \(e\) is safe then probe it w.p. \(\frac{y_e}{\alpha}\).
**A Simple 8-Approximation**

Analysis:

**Lemma:** For any edge \((u,v)\), at the point when \((u,v)\) is considered under \(\pi\), \(Pr(u \text{ loses its patience}) \leq 1/2\alpha\).

**Proof:** Let \(U\) be \#probes incident to \(u\) and before \(e\).

\[
\mathbb{E}[U] = \sum_{e \in \partial(u)} \Pr[\text{edge } e \text{ appears before } (u,v) \text{ in } \pi \text{ AND } e \text{ is probed}],
\]

\[
\cdot \sum_{e \in \partial(u)} \Pr[\text{edge } e \text{ appears before } (u,v) \text{ in } \pi] \cdot \frac{y_e}{\alpha},
\]

\[
= \sum_{e \in \partial(u)} \frac{y_e}{2\alpha} \cdot \frac{t_u}{2\alpha}.
\]

By the Markov inequality: \(Pr[U \geq t_u] \cdot \frac{\mathbb{E}[U]}{t_u} \cdot \frac{1}{2\alpha} \).
A SIMPLE 8-APPROXIMATION

Analysis:

Lemma: For any edge \(e=(u,v)\), at the point when \((u,v)\) is considered under \(\pi\), \(Pr(u \text{ is matched}) \leq 1/2\alpha\).

Theorem: The algorithm is a 8-approximation.

Proof: When \(e\) is considered,
\[
Pr(e \text{ is not safe}) \leq Pr(u \text{ is matched}) + Pr(u \text{ loses its patience}) + Pr(v \text{ is matched}) + Pr(v \text{ loses its patience}) \leq 2/\alpha
\]
Therefore,
\[
E[our \text{ solution}] = \sum_e w_e \Pr(e \text{ is safe}) \frac{y_e}{\alpha} p_e
\geq (1-2/\alpha) \frac{1}{\alpha} \sum_e w_e y_e p_e \geq 1/8 \text{ OPT} \quad (\alpha=4)
\]
Recall \(\sum_e w_e y_e p_e\) is an upper bound of \(\text{OPT}\).
AN IMPROVED APPROX. — BIPARTITE GRAPHS

Algorithm:
- \((x,y) \leftarrow \text{Optimal solution of the LP.}\)
- \(y' \leftarrow \text{Round } y \text{ to an integral solution using dependent rounding}\)
  [Gandhi et al. 06] and Let \(E' = \{e \mid y'_e = 1\}\).
  - (Marginal distribution) \(\Pr(y'_e = 1) = y_e\)
  - (Degree preservation) \(\text{Deg}_{E'}(v) \leq t_v\); (Recall \(\sum_{e \in \partial(\sigma)} y_e \leq t_v\))
  - (Negative Correlation) \(\Pr(\Lambda_{e \in S} (y'_e = 1)) \leq \prod_{e \in S} y_e\).
- Probe the edges in \(E'\) in random order.

For matching-probe model:
- \(M_1, \ldots, M_h \leftarrow \text{Optimal edge coloring of } E'\).
- Probe \(\{M_1, \ldots, M_h\}\) in random order.
FINAL REMARKS AND OPEN QUESTIONS

- Quite recently, Adamczyk has proved that the greedy algorithm is a 2-approximation for the unweighted version.

- Better approximations? (Unweighted: 2; Weighted bipartite: 3; Weighted: 4).

- $o(k)$-approximation for stochastic $k$-set packing? Or $\Theta(k)$ is the best possible?

- Any lower bound?
THANKS
AN IMPROVED APPROX. — BIPARTITE GRAPHS

Analysis (sketch):
Assume we have chosen $E'$.
Consider a particular edge $e= (u,v)$.
Let $B(e, \pi)$ be the set of incident edges that appear before $e$ in the random order $\pi$.

We claim that

$$\Pr [e \text{ is safe} \mid E'] \geq \mathbb{E}_{\pi} \left[ \prod_{f \in B(e, \pi)} (1 - p_f) \mid E' \right];$$

We claim that

$$\mathbb{E}_{\sigma} \left[ \prod_{f \in B(e, \sigma)} (1 - p_f) \mid E' \right] = \int_0^1 \prod_{f \in \partial_{E'}(e)} (1 - xp_f) \, dx.$$
AN IMPROVED APPROX. — BIPARTITE GRAPHS

Analysis cont: To see \( \mathbb{E}_\sigma \left[ \prod_{f \in B(e, \sigma)} (1 - p_f) \mid E' \right] = \int_0^1 \prod_{f \in \partial_{E'}(e)} (1 - xp_f) \, dx. \)

Consider this random experiment: For each edge in \( \partial_{E'}(e) \), we pick a random real in \([0,1]\). This produces a uniformly random ordering. Let r.v. \( A_f = (1-p_f) \) if \( f \) goes before \( e \), and \( A_f = 1 \) o.w.

Then, we consider \( \mathbb{E} \left[ \prod_{f \in \partial_{E'}(e)} A_f \right] = \int_0^1 \mathbb{E} \left[ \prod_{f \in \partial_{E'}(e)} A_f \mid a_e = x \right] \, dx \).
AN IMPROVED APPROX. — BIPARTITE GRAPHS

Analysis cont:

Define $\rho(r, p_{\text{max}})$ to be the optimal value of

$$\text{maximize } \int_0^1 \prod_{f \in \partial_{E'}(e)} (1 - x p_f) \, dx$$

subject to

$$\sum_{f \in \partial_{E'}(e)} p_f \cdot \ r$$

$$0 \cdot p_f \cdot p_{\text{max}}.$$ 

We can show $\rho(r, p_{\text{max}})$ is convex and decreasing on $r$. 
AN IMPROVED APPROX. — BIPARTITE GRAPHS

Analysis cont:

$$E[\text{ALG}] = \sum w_e p_e \Pr[e \in E'] \cdot \Pr[e \text{ was safe} \mid e \in \hat{E}]$$

Marginal Prob.
AN IMPROVED APPROX. — GENERAL GRAPHS

Algorithm:
- \((x,y) \leftarrow\) Optimal solution of the LP.
- Randomly partition vertices into \(A\) and \(B\).
- Run the previous algorithm on the bipartite graph \(G(A,B)\).

Thm: It is a \(2/\rho(1, p_{\text{max}})\)-approximation.
If \(p_{\text{max}} \to 1\), the ratio tends to 4. If \(p_{\text{max}} \to 0\), the ratio tends to \(2/(1-1/e) \approx 3.15\)