Constrained Matching Problem in Bipartite Graphs

Monaldo Mastrolilli  Georgios Stamoulis

ISCO '12

George Zirdelis, NTUA

June 28, 2013
Definition 1

Given a graph \( G = (V, E) \), a matching \( M \) in \( G \) is a set of pairwise non-adjacent edges; that is, no two edges share a common vertex.

Definition 2 (Bounded Color Matching)

Input is:
- Bipartite graph \( G(V, E) \) with bipartition \( V = V_1 \cup V_2 \).
- The edge set \( E \) is partitioned into \( k \) sets, \( E_1 \cup E_2 \cup \cdots \cup E_k \).
- Each edge set is characterized by a color \( j \in [k] \).
- Each edge \( e \in E \) has a profit \( p_e \in \mathbb{Q}^+ \).

Objective is:
- Find a maximum weight matching \( M \).
- In \( M \) there are no more that \( w_j \) edges of color \( j \) where \( w_j \in \mathbb{Z}^+ \), i.e.
  \( M \cap E_j \leq w_j, \quad \forall j \in [k] \).
LP of Bounded Color Matching

The relaxation of the IP for the Bounded Color Matching problem:

\[
\begin{align*}
\text{maximize} & \quad \max p^T x \\
\text{subject to} & \quad \sum_{e \in \delta(v)} x_e \leq 1, \quad \forall v \in V \\
& \quad \sum_{e \in E_j} x_e \leq w_j, \quad \forall j \in [k] \\
& \quad 0 \leq x_e \leq 1
\end{align*}
\]

where \(\delta(v)\) is the set of edges with one endpoints in \(v\). Integrality Gap is \(\frac{1}{2}\) so we cannot hope to achieve a better that \(\frac{1}{2}\) approximation algorithm.
Another way to describe the problem, say $\mathcal{M}'$, is the following:

$$\mathcal{M}' = \left\{ y \in \{0, 1\}^{\left| E \right|} : y \in \mathcal{M} \land \sum_{e \in E_j} y_e \leq w_j, \forall j \in [k] \right\}$$

where $\mathcal{M}$ is the usual bipartite polytope. Again, we can relax this by setting $y_e \in [0, 1]$. 
The Bounded Color Matching problem is known to be \textbf{NP – Complete}

Even if \( |E_j| \leq 2 \) and \( w_j = 1, \forall j \).

The special case of a 2-regular bipartite graphs where,

1. Each color appears twice.
2. Find a maximum matching with at most one edge per color

is \textbf{APX – Hard} so a \textbf{PTAS} is out of reach.
**Definition 3**

Let $E' \subseteq E$. Then we define the **characteristic vector** of $E'$ to be the binary vector $\chi_{E'} \in \{0, 1\}^{|E'|}$, s.t.

$$\chi_{E'}(e) = 1 \iff e \in E'$$

**Definition 4**

Let $y \in \mathbb{R}^n$. Then,

$$\text{support}(y) = \{i \in [n] : y_i \neq 0\}$$

i.e. the indices of the non-zero components of $y$. 

George Zirdelis, NTUA
Constrained Matching Problem in Bipartite Graphs
June 28, 2013
Lemma 5

Let $x^*$ be an optimal basic feasible solution for the LP described by $\mathcal{M}'$ s.t. $x_e^* > 0$, $\forall e \in E$. Then, there exist $F \subseteq V$ and $Q \subseteq [k]$ s.t.,

1. $\sum_{e \in \delta(v)} x_e^* = 1$, $\forall v \in F$.
2. $\sum_{e \in E_j} x_e^* = w_j$, $\forall j \in Q$.
3. $\{\chi_{\delta(v)}\}_{v \in F}$ and $\{\chi_{E_j}\}_{j \in Q}$ are all linearly independent.
4. $|E| = |F| + |Q|$ where $|E|$ is the number of the edges with $x_e^* > 0$.

- If $\sum_{e \in \delta(v)} x_e = 1$ then $v$ is a tight vertex.
- If $\sum_{e \in E_j} x_e = w_j$ then $j$ is a tight color class.
Define the **residual graph** to be the graph with the same vertex set but we include an edge \( e \) if \( x_e > 0 \) in the LP solution for the original graph.

**Lemma 6**

*Take any basic feasible solution \( x \) s.t. \( x_e > 0, \forall e \), i.e. we remove any edge with \( x_e = 0 \). Then one of the following must be true:*

1. **either there is an edge s.t.** \( x_e = 1 \).
2. **or there is a color class** \( j \in Q \subseteq [k] \) s.t. \( |E_j| \leq w_j + 1 \) **in the residual graph**
3. **or there is a tight vertex** \( v \in F \) s.t. **the degree of** \( v \) **is 2 in the residual graph.**

Using the lemma above will can do iterative rounding to obtain a solution.
$C, E$ will be resp. the set of the available colors and edges, at each round.

Initialize $M = \emptyset$.
While $C \neq \emptyset$ or $E \neq \emptyset$ do:

1. Compute an optimal (fractional) basic solution $x$ to the current LP.
2. Remove all edges from the graph s.t. $x_e = 0$.
3. Remove all vertices of the graph s.t. $\deg(v) = 0$.
4. if $\exists e = (u, v) \in E: x_e = 1$ and $e \in C_j$ then $M := M \cup \{e\}$, $V = V \setminus \{u, v\}$, $w_j := w_j - 1$. if $w_j = 0$ then $C := C \setminus C_j$, $E := E \setminus \{e : e \in E_j\}$.
5. (Relaxation:) while $V \cup C \neq \emptyset$
   - if $\exists$ color class $C_j \in Q$ with $|E_j| \leq w_j + 1$ then remove the constraint for this color class, i.e. define $C := C \setminus C_j$.
   - if $\exists$ vertex $v \in F$ s.t. $\deg(v) = 2$ then remove the constraint for that vertex.

Return $M$

At each step of the algorithm, either we add an edge to our matching $M$, or we remove a tight constraint. Thus the algorithm will terminate in at most $|Q| + |F|$.
Since we remove the degree constraints for a vertex $v$ when $\text{deg}(v) = 2$ we select edges from a graph $G'$ that is a collection of disjoint paths or cycles.

But a disjoint path or cycle can be partitioned into two matchings, i.e. $M_1, M_2$ and we select the one with the highest profit, i.e.

$$\max(p(M_1), p(M_2)) \geq \frac{1}{2}p(M_1 \cup M_2)$$

Therefore we do this for every connected component (disjoint paths and cycles), we get at least $\frac{1}{2}$ of the profit of the matchings but we violate by an additive 1 every color constraint.

As a result of the above there is a polynomial time $(1/2, \text{additive 1})$ bi-criteria approximation algorithm for the weighted Bounded Color Matching problem.
We now consider the unweighted version of the Bounded Color Matching problem: Compute a maximum cardinality matching \( M \) s.t. in \( M \) we have at most \( w_j \) edges for color class \( j \).

Recall that from Lemma 7 we have that for any solution to the LP, if \( 0 < x_e < 1 \) then,

- either there exists a tight color class \( j \in Q \) s.t. \( |\text{support}(x) \cap E_j| \leq w_j + 1 \)
- or there exists a tight vertex \( v \in F \) s.t. \( \text{deg}(v) = 2 \).
The main idea of the algorithm for the cardinality version consists of the two following steps:

- **Relaxation step**: We identify a tight color class $j$ and we remove its constraint, thus relaxing the problem.
- **Rounding step**
  - We round appropriately some variables to 1 and some others to 0, preserving feasibility.
  - Rounding step comes with a parameter $\lambda \in [0, 1]$. Idea is that if we round $x_e$ to 1, we need to update the color bound of this color class.
  - Using $\lambda$ we update the color bound by any value in $[x_e, 1]$ (if we use $x_e + \lambda(1 - x_e)$).
  - Values of $\lambda$ closer to $x_e$ violate more the color constraint whereas values closer to 1 give less violation but worst performance guarantee.
Lemma 7

Let $x$ be the optimal solution in $G$ (as stated in Lemma 4) before the rounding step and $\hat{x}$ be the optimal solution after the rounding step in $\hat{G}$. Then we have that,

$$\sum_{e \in E(G)} x_e - \sum_{e \in E(\hat{G})} \hat{x}_e \leq 1 + (\gamma + \lambda \gamma)$$

where $\gamma = 1 - x_e$.

The loss due to a single rounding step is at most $\gamma + \lambda \gamma$ which can be at most $\frac{1}{2}(\lambda + 1)$. 

$C, E$ will be resp. the set of the available colors and edges, at each round.

Initialize $M = \emptyset$.
While $C \neq \emptyset$ or $E \neq \emptyset$ do:

1. Compute an optimal (fractional) basic solution $x$ to the current LP.
2. Remove all edges from the graph s.t. $x_e = 0$.
3. Remove all vertices of the graph s.t. $\deg(v) = 0$.
4. if $\exists e = (u, v) \in E : x_e = 1$ and $e \in C_j$ then $M := M \cup \{e\}$, $V = V \setminus \{u, v\}$, $w_j := w_j - 1$. if $w_j = 0$ then $C := C \setminus C_j$, $E := E \setminus \{e : e \in E_j\}$.
5. (Relaxation:) If $\exists$ color class $j \in Q$ with $|E| \leq \lceil w_j \rceil + 1$ then remove the constraint for this color class, i.e. set $C := C \setminus C_j$ and iterate.
6. (Rounding:) if $\exists v \in F$ s.t. $\deg(v) = 2$ then let: $u_1, u_2$ be the neighbors of $v$ and let $e_1, e_2$ be the two edges incident on $v$. Assume w.l.o.g. that $x_{e_1} \geq \frac{1}{2}$ and $e_1 = (u_1, v)$.
   - Round $x_{e_1}$ to 1. Add it $(e_1)$ to $M$.
   - Round $x_{e_2}$ and all other edges incident to $u_1$ to zero.
   - If $e_1 \in E_j$ then set $w_j := w_j - x_{e_1} - \lambda(1 - x_{e_1})$.
   - Remove $v, u_1$ and all the rounded edges from the graph and iterate.

Return $M$
From Lemma 7 we have that in each rounding step the objective function decreases by $1 + \gamma + \lambda \gamma$.

Intuitively, the larger the value of $\gamma$ is, the fewer iterations the algorithm will perform.

Because $OPT \leq |V|/2$ and at each rounding step we delete 2 vertices from the current graph, we can perform at most $|V|/4$ rounding steps. So, we can have at most $|V|/4$ values of $\gamma$, though they all might be different.
Lemma 8

Let $\tilde{x}$ be the final solution of the algorithm that corresponds to $M$. Then we have that,

$$\sum_{e \in M} \tilde{x}_e \geq \frac{2}{3 + \lambda} \sum_{e \in E(G)} x_e$$

Proof.

- Since we choose $x_{e_1} \geq 1/2$ assume that in some iteration $\gamma_1 = \frac{p}{q} \in (0, 1/2]$ and also that this $\gamma_1$ appears $k_1$ times during the Rounding steps.

- The total decrease in the objective function is $\frac{q+p(\lambda+1)}{q} = 1 + \gamma_1 + \lambda \gamma_1$

- Maximum number of iterations we can have for this particular $\gamma_1$ is $OPT \cdot \frac{q+p(\lambda+1)}{q}$ before it truncates to 0. E.g. for $\gamma_1 = \frac{1}{3}$ and $\lambda = \frac{1}{2}$ in the next iteration of the LP we will have

  $$OPT' = OPT - \frac{3}{2} \Rightarrow OPT - OPT' = \frac{3}{2}$$

and so we can have at most $OPT \cdot \frac{2}{3}$ iterations.
Assume that the algorithm performs a fraction $f_i$ of the maximum possible number of iterations for each $\gamma_i$. Then we have that,

$$\sum_i f_i \leq 1$$

because in each round we reduce the objective function.

At the end of the algorithm the final objective function value will be,

$$OPT - \sum_i f_i \cdot \frac{OPT}{1 + \gamma_i + \lambda \gamma_i} \cdot \gamma_i(\lambda + 1)$$
Set \( g(\gamma_i) = \frac{\gamma_i(\lambda+1)}{1+\gamma_i+\lambda\gamma_i} \) which monotonically increases. We have that

\[
SOL = OPT - OPT \sum_i f_i \cdot g(\gamma_i) \geq OPT - OPT \sum_i f_i \cdot g(1/2)
\]

\[
= OPT - OPT \sum_i \frac{\lambda + 1}{\lambda + 3}
\]

\[
\geq \frac{2}{\lambda + 3} OPT
\]

Using similar arguments one can show that the color bound \( w_j \) of a color \( j \) can be violated by at most a factor of \( \frac{2}{1+\lambda} w_j + 1 \). 

\( \square \)
Theorem 9

For any $\lambda \in [0, 1]$, there is a polynomial time $(\frac{2}{3+\lambda}, \frac{2}{1+\lambda}w_j + 1)$ bi-criteria approximation algorithm for the Bounded Color Matching problem.

- The closer $\lambda$ is to 1 the more we deteriorate from the optimal objective function value but the less we lose in color bounds.
- The closer $\lambda$ is to 0 the more we violate the color constraints but the better the approximation guarantee is.
- Depending on the application we choose a parameter $\lambda$ that is more suitable.
- We have a family of algorithms for the unweighted case.
Monaldo Mastrolilli and Georgios Stamoulis.

Constrained matching problems in bipartite graphs.