Approximation Algorithms for Finding Maximum Independent Sets in Unions of Perfect Graphs

Authors: Chakaravarthy, Pandit, Roy, Sabharwal
Speaker: Arindam Pal

Department of Computer Science and Engineering
Indian Institute of Technology Delhi

November 30, 2012

Agenda

- The Maximum Independent Set Problem in Graphs
- Perfect Graphs and their Properties
- Summary of Results
- The Common Independent Set Problem
- LP Formulation for 2-$\textsc{MaxCIS}$
- The Rounding Algorithm and Analysis

The Maximum Independent Set Problem in Graphs

- Given a graph $G = (V, E)$, and a profit (weight) function $p : V \mapsto \mathbb{R}^+$, the goal is to find an independent set of maximum profit.
- If $p(v) = 1, \forall v \in V$, this is the cardinality version.
- [ZUCKERMAN]: It is NP-Hard to approximate MIS within a factor of $n^{1-\epsilon}$, for any $\epsilon > 0$.
- However, there are easy families of graphs for which MIS can be solved optimally in polynomial time.
- Examples are $\{$*interval, chordal, comparability*$\}$ graphs.
- All of them are *perfect graphs*.

What is a Perfect Graph?

- A perfect graph is one in which the chromatic number of every induced subgraph equals the size of the largest clique of that subgraph. For every $H \subseteq G, \chi(H) = \omega(H)$.
- The Perfect Graph Theorem [Lovasz]: A graph is perfect if and only if its complement graph is also perfect.
- The Strong Perfect Graph Theorem [Chudnovsky, Robertson, Seymour, and Thomas]: A graph is perfect if and only if it has no induced subgraph that is an odd cycle of length at least five or its complement. These graphs are called *Berge graphs*.
- Perfect graphs can be recognized in polynomial time.
- For perfect graphs, graph coloring, maximum clique, and maximum independent set problems can all be solved in polynomial time using the ellipsoid method.

Common Independent Set

- A subset of vertices $X \subseteq V$ is called a common independent set (CIS) of graphs $G_i = (V, E_i), 1 \leq i \leq t$, if X is an independent set in each graph G_i.
- Alternatively, a CIS is an independent set in the union graph $G = (V, E)$, where $E = \bigcup_{i=1}^{t} E_i$.
- The goal is to find the maximum weight CIS ($\textsc{MaxCIS}$).
- For a fixed constant k, the k-$\textsc{MaxCIS}$ problem is the special case where the number of input graphs is $t = k$.
- We consider restricted versions of the MaxCIS problem, where all the input graphs belong to a particular class of graphs $\mathcal{C}$.
- In this talk, $\mathcal{C}$ will be the family of perfect graphs.

Summary of Results

- An $O(\sqrt{n})$-approximation algorithm for the weighted 2-$\textsc{MaxCIS}$ problem on perfect graphs.
- Can be easily generalized to an $O(n^{\frac{k-1}{k}})$-approximation algorithm for the weighted k-$\textsc{MaxCIS}$ problem on perfect graphs.
- The LP has an integrality gap of $\sqrt{n}$, even when both the graphs are comparability graphs and all the weights are unit.
- A stronger LP has an integrality gap of $n^{0.16}$, for unweighted comparability graphs.
- If $\mathrm{P} \neq \mathrm{NP}$, then the 2-$\textsc{MaxCIS}$ problem on comparability graphs cannot be approximated within any constant factor.
- If $\mathrm{NP} \not\subseteq \mathrm{DTIME}[n^{O(\log n)}]$, then 2-$\textsc{MaxCIS}$ on comparability graphs cannot be approximated within a factor of $2^{\sqrt{\log n}}$.

Weighted 2-MaxCIS Problem on Perfect Graphs

- Write an (exponential size) LP relaxation.
- Compute an optimal fractional solution for the LP.
- Partition the vertex set into sets of large and small vertices.
- Recover an $O(\sqrt{n})$ fraction of the LP solution for both large and small instances.
- This gives an $O(\sqrt{n})$-approximation algorithm.

Linear Programming Formulation

We give a natural linear programming formulation for 2-$\textsc{MaxCIS}$. Here $x(v)$ denotes whether the vertex $v \in V$ is included in the independent set. $\mathcal{C}_i$ is the set of all cliques in G_i.

$$\begin{aligned} \textbf{maximize} \quad & \sum_{v \in V} p(v) x(v) \\ \textbf{such that} \quad & \sum_{v \in C} x(v) \leq 1 \qquad \forall C \in \mathcal{C}_1 \cup \mathcal{C}_2 \\ & 0 \leq x(v) \leq 1 \qquad \forall v \in V \end{aligned}$$

This LP has exponentially many constraints in n. How to solve it?

Separation Oracle

- Assign $x(v)$ as the profit of v.
- Find the maximum profit clique in G_1 and G_2.
- If both these cliques have profit at most 1, then x is a feasible solution.
- Otherwise, the constraint corresponding to the clique having profit greater than 1 is violated.

A Preparatory Lemma

Lemma

For any $X \subseteq V$, there exists a $I \subseteq X$ such that I is an independent set in $G = G_1 \cup G_2$ and $p(I) \geq \frac{p(X)}{\omega_1(X)\omega_2(X)}$.

- G_1 is perfect, so $G_1[X]$ can be colored with $\omega_1(X)$ colors.
- So, X can be partitioned into $\omega_1(X)$ color classes (independent sets), one of which must have profit at least $\frac{p(X)}{\omega_1(X)}$.
- Compute the maximum profit independent set I_1 in $G_1[X]$.
- Clearly, $p(I_1) \geq \frac{p(X)}{\omega_1(X)}$.

Proof of Lemma ...

- $G_2[I_1]$ can be colored with $\omega_2(I_1) \leq \omega_2(X)$ colors.
- Hence, $G_2[I_1]$ has an independent set of profit at least $\frac{p(I_1)}{\omega_2(X)}$.
- Compute the maximum profit independent set I in $G_2[I_1]$.
- Note that, $p(I) \geq \frac{p(I_1)}{\omega_2(X)} \geq \frac{p(X)}{\omega_1(X)\omega_2(X)}$.
- Moreover, I is an independent set in both G_1 and G_2. So, it is an independent set in G.

The Rounding Algorithm

- Let x be the optimal fractional solution to the LP.
- Let $LP^\star(X) = \sum_{v \in X} p(v)x(v), X \subseteq V$.
- Partition the vertex set V in two parts, SML and LRG, the set of small and large vertices respectively.
- $SML = \left\{v : 0 \leq x(v) \leq \frac{1}{\sqrt{n}}\right\}$.
- $LRG = \left\{v : \frac{1}{\sqrt{n}} < x(v) \leq 1\right\}$.
- $LP^\star = LP^\star(V) = LP^\star(SML) + LP^\star(LRG)$ is the profit of the optimal LP solution.
- Let $P_{\max} = \max_{v \in V} p(v)$ be the maximum profit and let $v_{\max}$ be the corresponding vertex.

Handling Large Vertices

- Define, $U_i = \left\{ v : \frac{2^i}{\sqrt{n}} < x(v) \leq \frac{2^{i+1}}{\sqrt{n}} \right\}$, for $0 \leq i \leq \ell - 1$, where $\ell = \frac{1}{2} \log n$.
- $U_0, \ldots, U_{\ell-1}$ forms a partition of LRG.
- Using the Lemma on U_j, find an independent set I_j of G such that $p(I_j) \geq \frac{p(U_j)}{\omega_1(U_j)\omega_2(U_j)}$, for $0 \leq j \leq \ell - 1$.
- Among these ℓ independent sets, let $I^\star$ be the one having the maximum profit.
- If $p(I^\star) > P_{\max}$, output $I^\star$; else output $\{v_{\max}\}$.
- Let I be the independent set output by the algorithm.

Analysis for Large Vertices

Lemma

$$LP^{\star}(U_j) \leq \left(\frac{2\sqrt{n}}{2^j}\right) \cdot p(I_j), 0 \leq j \leq \ell - 1.$$

- Define $\beta = \frac{2^j}{\sqrt{n}}$. We have to show that $LP^{\star}(U_j) \leq \frac{2p(I_j)}{\beta}$.
- We know that, $\beta \leq x(v) \leq 2\beta$.
- Hence, $LP^{\star}(U_j) \leq 2\beta \cdot p(U_j)$.
- Let $\omega_{\min} = \min\{\omega_1(U_j), \omega_2(U_j)\}$ and $\omega_{\max} = \max\{\omega_1(U_j), \omega_2(U_j)\}$.

- There exists a subset $C \subseteq U_j$, which is a clique in $G_1[X]$ or $G_2[X]$ such that $|C| = \omega_{\max}$.
- The LP contains a constraint corresponding to this clique C.
- Hence, $\sum_{v \in C} x(v) \leq 1$.
- Since $x(v) \geq \beta$, we have $\beta\omega_{\max} \leq 1$, i.e., $\beta \leq \frac{1}{\omega_{\max}}$.
- Thus, $LP^{\star}(U_j) \leq \frac{2p(U_j)}{\omega_{\max}}$.
- Since, $p(I_j) \geq \frac{p(U_j)}{\omega_{\min}\omega_{\max}}$, it follows that $LP^{\star}(U_j) \leq 2\omega_{\min}p(I_j)$.
- But, $\omega_{\min} \leq \omega_{\max} \leq \frac{1}{\beta}$. So, $LP^{\star}(U_j) \leq \frac{2p(I_j)}{\beta}$.

Lemma

$$LP^\star(LRG) \le 4\sqrt{n} \cdot p(I^\star).$$

Proof.

$$\begin{aligned} LP^\star(LRG) &= \sum_{j=0}^{\ell-1} LP^\star(U_j) \\ &\le \sum_{j=0}^{\ell-1} \left(\frac{2\sqrt{n}}{2^j}\right) \cdot p(I_j) \\ &\le 2\sqrt{n} \cdot p(I^\star) \sum_{j=0}^{\ell-1} \left(\frac{1}{2^j}\right) \\ &\le 4\sqrt{n} \cdot p(I^\star). \end{aligned}$$

□

The Final Analysis

Lemma

$$LP^\star \leq 5\sqrt{n} \cdot p(I).$$

- Since $p(I) \geq P_{\max}, LP^\star(SML) \leq \sqrt{n} \cdot P_{\max} \leq \sqrt{n} \cdot p(I)$.
- Moreover, $LP^\star(LRG) \leq 4\sqrt{n} \cdot p(I^\star) \leq 4\sqrt{n} \cdot p(I)$.
- Hence, $LP^\star = LP^\star(SML) + LP^\star(LRG) \leq 5\sqrt{n} \cdot p(I)$.
- Therefore, this is an $O(\sqrt{n})$-approximation algorithm.

An Interesting Observation

- In perfect graphs, MIS can be computed in polynomial time.
- In union of k perfect graphs, MIS can be approximated within a factor of $O(n^{\frac{k-1}{k}})$.
- As $k \to \infty$, the approximation factor approaches $O(n)$, which is no better than general graphs.
- This indicates that as we add more perfect graphs in the union, the *perfection* of the resulting graph reduces.