New insight into introducing a $(2 - \varepsilon)$-approximation ratio for minimum vertex cover problem

Majid Zohrehbandian

Department of Mathematics, Karaj Branch, Islamic Azad University, Karaj, Iran.

zohrebandian@yahoo.com

Abstract

Vertex cover problem is a famous combinatorial problem, which its complexity has been heavily studied over the years. It is known that it is hard to approximate to within any constant factor better than 2, while a 2-approximation for it can be trivially obtained. In this paper, new properties and new techniques are introduced which lead to approximation ratios smaller than 2 on special graphs. Then, by a combination of semidefinite programming and a rounding procedure, along with satisfying the proposed assumptions, we introduce an approximation algorithm with a performance ratio of 1.999999 on arbitrary graphs.

Keywords: Discrete Optimization, Vertex Cover Problem, Complexity Theory, NP-Complete Problems.

MSC 2010: 90C35**,** 90C60.

1. Introduction

In complexity theory, the abbreviation NP refers to "nondeterministic polynomial", where a problem is in NP if we can quickly (in polynomial time) test whether a solution is correct. P and NP-complete problems are subsets of NP Problems. We can solve P problems in polynomial time while determining whether or not it is possible to solve NP-complete problems quickly (called the P vs NP problem) is one of the principal unsolved problems in Mathematics and Computer science.

Here, we consider the vertex cover problem which is a famous NP-complete problem. It cannot be approximated within a factor of 1.36 [1], unless $P = NP$, while a 2-approximation factor for it can be trivially obtained by taking all the vertices of a maximal matching in the graph. However, improving this simple 2-approximation algorithm has been a quite hard task [2,3].

In this paper, we introduce a $(2 - \varepsilon)$-approximation ratio on special graphs, and then, we show that on arbitrary graphs a $(2 - \varepsilon)$-approximation ratio can be obtained by a combination of semidefinite programming (SDP) and a rounding procedure. The rest of the paper is structured as follows. Section 2 is about the vertex cover problem and introduces new properties and new techniques which lead to a $(2-\varepsilon)$-approximation ratio on special graphs. In section 3, we propose a rounding procedure along with using the satisfying properties to propose an algorithm with a performance ratio smaller than 2 on arbitrary graphs. Finally, Section 4 concludes the paper.

2. Introducing a $(2-\varepsilon)$-approximation ratio on special graphs

In the mathematical discipline of graph theory, a vertex cover of a graph is a set of vertices such that each edge of the graph is incident to at least one vertex of the set. The problem of finding a minimum vertex cover is a typical example of an NP-complete optimization problem. In this section, new properties and new techniques are introduced which lead to approximation ratios smaller than 2 on special problems.

Let $G=(V,E)$ be an undirected graph on vertex set V and edge set E, where $|V|=n$. Throughout this paper, suppose that the vertex cover problem on G is hard and we have produced an arbitrary feasible solution for the problem, with vertex partitioning $V=V_{1G}\cup V_{-1G}$ (V_{1G} is a vertex cover of the graph G) and objective value $|V_{1G}|$.

By defining the decision variables x_j and x_{ij} as follows:

$$x_j=\begin{cases}+1 & j\in V_{1G}^*\\ -1 & j\in V_{-1G}^*\end{cases}$$

$$x_{ij}=\begin{cases}+1 & i,j\in V_{1G}^* \quad or \quad i,j\in V_{-1G}^*\\ -1 & otherwise\end{cases}$$

We can introduce the following integer linear programming (ILP) model for the minimum vertex cover problem:

$$(1)\qquad \min_{s.t.} \quad z^1=\sum_{1\le j\le n}\frac{1+x_j}{2}$$

$$\begin{array}{ll}
+x_i+x_j-x_{ij}=+1 & ij\in E, 1\le i<j\le n\\
+x_{ij}+x_{jk}+x_{ik}\ge -1 & 1\le i<j<k\le n\\
+x_{ij}-x_{jk}-x_{ik}\ge -1 & 1\le i<j<k\le n\\
-x_{ij}+x_{jk}-x_{ik}\ge -1 & 1\le i<j<k\le n\\
-x_{ij}-x_{jk}+x_{ik}\ge -1 & 1\le i<j<k\le n\\
x_j, x_{ij}\in\{-1,1\} & 1\le i<j\le n
\end{array}$$

Moreover, by consideration of x_j's as x_{oj} and addition of the constraint $x\succcurlyeq 0$, we have the well known SDP formulation as follows:

$$(2) \qquad \min_{s.t.} \quad z^2 = \sum_{i \in V} \frac{1 + v_o v_i}{2}$$

$$+v_o v_i + v_o v_j - v_i v_j = 1 \qquad ij \in E$$

$$+v_i v_j + v_i v_k + v_j v_k \geq -1 \qquad i, j, k \in V \cup \{o\}$$

$$+v_i v_j - v_i v_k - v_j v_k \geq -1 \qquad i, j, k \in V \cup \{o\}$$

$$-v_i v_j + v_i v_k - v_j v_k \geq -1 \qquad i, j, k \in V \cup \{o\}$$

$$-v_i v_j - v_i v_k + v_j v_k \geq -1 \qquad i, j, k \in V \cup \{o\}$$

$$v_i v_i = 1 \qquad i \in V \cup \{o\}$$

$$v_i v_j \in \{-1, +1\} \qquad i, j \in V \cup \{o\}$$

Theorem 1. Suppose that $z^{2*} \geq \frac{n}{2} + \frac{n}{k} = \frac{(k+2)n}{2k}$. Then, for all feasible solutions $V = V_{1G} \cup V_{-1G}$ we have the approximation ratio $\frac{|V_{1G}|}{Z^{2*}} \leq \frac{2k}{k+2}$.

Proof. $\frac{|V_{1G}|}{Z^{2*}} \leq \frac{n}{Z^{2*}} \leq \frac{2k}{k+2} < 2$ ■

Assumption 1. From now on, we assume that $\frac{n}{2} \leq z^{2*} < \frac{n}{2} + \frac{2n}{1000000}$; Otherwise for all feasible solutions $V = V_{1G} \cup V_{-1G}$ we have the approximation ratio $\frac{|V_{1G}|}{Z^{2*}} \leq \frac{2 \times \frac{1000000}{2}}{\frac{1000000}{2} + 2} < 1.999993 < 2.$

Theorem 2. Suppose that we have a suitable feasible solution $V_{1G} \cup V_{-1G}$ for which we have $|V_{1G}| \leq k|V_{-1G}|$. Then, we have an approximation ratio $\frac{|V_{1G}|}{Z^{2*}} \leq \frac{2k}{k+1} < 2$.

Proof. $\exists t \leq k$, for which we have $|V_{1G}| = t|V_{-1G}| = t\frac{n}{t+1}$. Then, $z^{2*} \geq \frac{n}{2} = \frac{t+1}{2t}|V_{1G}|$ which concludes that $\frac{|V_{1G}|}{Z^{2*}} \leq \frac{2t}{t+1} \leq \frac{2k}{k+1}$ ■

Therefore, for bounded values of k, we have some approximation ratios smaller than 2. But, if $k \to \infty$ then $\frac{|V_{1G}|}{Z^{2*}} \to 2$.

Corollary 1. Suppose that we have a suitable feasible solution $V_{1G} \cup V_{-1G}$. If $|V_{1G}| < \frac{k}{k+1}n$ then $|V_{1G}| < k|V_{-1G}|$ and $\frac{|V_{1G}|}{Z^{2*}} < \frac{2k}{k+1} < 2$.

Assumption 2. We don't have a suitable feasible solution $V = V_1 \cup V_{-1}$ for which $|V_1| < \frac{999999}{1000000}n$; Otherwise, for this feasible solution we have an approximation ratio $\frac{|V_1|}{Z^{2*}} \leq \frac{2 \times 999999}{999999+1} = 1.999998 < 2$.

Up to now, we could introduce a $(2-\varepsilon)$-approximation ratio on special graphs with suitable characteristics. In section 3, we are going to introduce such a ratio on arbitrary graphs, where we assume that we have $V = V_1 \cup V_{-1}$ as a feasible solution of the vertex cover problem on arbitrary graph G for which $|V_1| \geq 0.999999n$ and $\frac{n}{2} \leq z^{2*} < \frac{n}{2} + \frac{2n}{1000000}$.

3. A (1.999999)-approximation algorithm for the vertex cover problem

In section 2, we could introduce a $(2-\varepsilon)$-approximation ratio on graphs without the proposed assumptions. Here, we are going to introduce a 1.999999-approximation ratio on arbitrary graphs. To do this, we assume the following assumption.

Assumption 3. By solving the SDP relaxation (2),

a) For less than $\frac{1}{1000000}n$ of vertices $j \in V$ and corresponding vectors we have $v_o^* v_j^* < 0$; Otherwise based on these vertices, we have a feasible solution with $|V_{-1}| \geq \frac{1}{1000000}n$, $|V_1| \leq \frac{999999}{1000000}n$ and an approximation ratio $\frac{|V_{1G}|}{Z^{2*}} < \frac{2(999999)}{999999+1} = 1.999998 < 2$.

b) For less than $\frac{1}{100}n$ of vertices $j \in V$ and corresponding vectors we have $v_o^* v_j^* > 0.0004$. Otherwise, $z^{2*} \geq \underbrace{\left(\frac{1+(-1)}{2} \times \frac{n}{1000000}\right)}_{v_o^* v_j^* < 0} + \underbrace{\left(\frac{1+0}{2} \times \frac{989999n}{1000000}\right)}_{0 \leq v_o^* v_j^* \leq 0.0004} + \underbrace{\left(\frac{1+0.0004}{2} \times \frac{n}{100}\right)}_{v_o^* v_j^* > 0.0004} = \frac{n}{2} + \frac{3n}{2000000}$ and for all feasible solutions, we have the approximation ratio $\frac{|V_{1G}|}{Z^{2*}} < \frac{2(\frac{2000000}{3})}{\frac{2000000}{3}+2} < 1.999995 < 2$.

Definition 1. Let $\varepsilon = 0.0004$ and $G_\varepsilon = \{j \in V | 0 \leq v_o^* v_j^* \leq +\varepsilon\}$.

Based on Assumption (3), for more than $\frac{989999}{1000000}n$ of vertices $j \in V$ and corresponding vectors we have $0 \leq v_o^* v_j^* \leq +\varepsilon$; i.e. $|G_\varepsilon| \geq 0.989999n$.

Theorem 3. For any normalized vector w, the induced subgraph on $H_w = \{j \in G_\varepsilon;\ |w v_j^*| > 0.5003\}$ is a bipartite graph.

Proof. Let us divide the vertex set H_w as follows:

$$S = \{j \in H_w |\ w v_j^* < -0.5003\} \quad \text{and} \quad T = \{j \in H_w |\ w v_j^* > +0.5003\}$$

Then, it is sufficient to show that the sets S and T are null subgraphs. For each edge $ij \in E(G)$ and based on the first constraint of the SDP model (2), if $i, j \in H_w \subseteq G_\varepsilon$ then we have $v_i^* v_j^* \leq -1 + 2\varepsilon$. Therefore, if $i, j \in T$ then the triangle inequality between vectors w, v_i^* and v_j^* is violated; i.e.

$$\|v_i^* - v_j^*\| \leq \|w - v_i^*\| + \|w - v_j^*\|$$

$$\sqrt{2-2v_i^*v_j^*} \leq \sqrt{2-2wv_i^*} + \sqrt{2-2wv_j^*}$$

$$\sqrt{2-2(-1+2(0.0004))} \leq \sqrt{2-2v_i^*v_j^*} \leq \sqrt{2-2wv_i^*} + \sqrt{2-2wv_j^*} \leq 2\sqrt{2-2(0.5003)}$$

Therefore, we have $1.9995999 \cong \sqrt{3.9984} \leq 2\sqrt{9994} \cong 1.9993999$, which is a contradiction.

Likewise, if $i, j \in S$ then the triangle inequality between vectors $-w, v_i^*$ and v_j^* is violated ■

Corollary 2. If $\exists k \in V: |H_k| \geq \frac{n}{1000000}$, where $H_k = \{j \in G_\varepsilon;\ |v_k^* v_j^*| > 0.5003\}$, then we have a feasible solution $V_{1G} \cup V_{-1G}$, correspondingly, where $|V_{-1G}| = max\{|S|, |T|\} \geq \frac{n}{2000000}$. Hence, $|V_{1G}| \leq 1999999|V_{-1G}|$ and we have $\frac{|V_{1G}|}{z^{2*}} \leq \frac{2\times1999999}{1999999+1} = 1.999999 < 2$.

Corollary 3. By introducing a normalized random vector w, where $|H_w| \geq \frac{n}{1000000}$, we have a feasible solution $V_{1G} \cup V_{-1G}$, correspondingly, where $|V_{-1G}| = max\{|S|, |T|\} \geq \frac{n}{2000000}$. Hence, we have $|V_{1G}| \leq 1999999|V_{-1G}|$ and $\frac{|V_{1G}|}{z^{2*}} \leq \frac{2\times1999999}{1999999+1} = 1.999999 < 2$.

Theorem 4. Let u, w be two normalized random vectors, then for any normalized vector v_j^*, we have $\Pr(|uv_j^*| \leq 0.5003\ \&\ |wv_j^*| \leq 0.5003) < 0.753$.

Proof. Let $v_j^* = v_j' + v_j''$, where v_j' is the projection of vector v_j^* onto the $u - w$ plane (suppose that the vector u is on the ox axis) and v_j'' is the projection of v_j^* onto the normal vector of that plane. Then, $|uv_j^*| = |uv_j'| \leq 0.5003$ if and only if the vector v_j' is projected on the gray region in the first quadrant (or its symmetric region with respect to the oy axis, the ox axis and the origin in the second, fourth and third region), where $|v_j'| \leq f(\theta) = \frac{0.5003}{\cos\theta}$. In this manner, $|v_j'||u|\cos\theta \leq 0.5003$; See Figure 1.

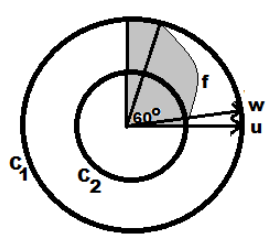

Figure 1. The $u - w$ plane.

Therefore, $\Pr(|uv_j^*| \leq 0.5003\ \&\ |wv_j^*| \leq 0.5003) \cong \frac{S}{2\pi}$, where S is the area of the common gray region between two vectors u and w. Note that, the maximum area of the region S and corresponding probability is produced based on the $|uw| \cong 1$ condition (and the minimum value for the probability is produced when $|uw| \cong 0$). But, $S = 4(\int_0^{\frac{\pi}{3}} f(\theta)d\theta + \int_{\frac{\pi}{3}}^{\frac{\pi}{2}} d\theta) = 4(\int_0^{\frac{\pi}{3}} \frac{0.5003}{\cos\theta} d\theta + \frac{\pi}{6}) \cong 4(1.182473)$. Hence, we have:

$\Pr(|uv_j^*| \leq 0.5003 \ \& \ |wv_j^*| \leq 0.5003) \cong \frac{4(1.182473)}{2\pi} = 0.752786 < 0.753$ ■

Therefore, by introducing two normalized random vectors u, w, for at most $0.753n$ of the vectors v_j^* (the optimal solution of the SDP model) we have $|uv_j^*| \leq 0.5003$ and $|wv_j^*| \leq 0.5003$, and therefore, for at least $0.247n$ of the vectors v_j^* we have $|uv_j^*| > 0.5003$ or $|wv_j^*| > 0.5003$.

Hence, one of the two bipartite graphs H_u or H_w has more than $\frac{0.247n}{2}$ of vertices which produces a null subgraph with more than $\frac{0.247n}{4} = 0.06175n$ of the vertices and based on the Corollary (3) we have an approximation ratio $\frac{|V_{1G}|}{z^{2*}} \leq 1.999999 < 2$.

Now, we can introduce our algorithm to produce an approximation ratio $\rho \leq 1.999999$.

Zohrehbandian Algorithm (To produce a vertex cover solution with a factor $\rho \leq 1.999999$)

Step 1. Solve the SDP (2) relaxation.

Step 2. If for more than $\frac{n}{1000000}$ of vertices $j \in V$ and corresponding vectors we have $v_o^* v_j^* < 0$, then produce the suitable solution $V_{1G} \cup V_{-1G}$, correspondingly, where $V_{-1G} = \{j | v_o^* v_j^* < 0\}$. Therefore, based on the Assumption (3. a) we have $\frac{|V_{1G}|}{z^{2*}} \leq 1.999999$. Otherwise, go to Step 3.

Step 3. If for more than $\frac{1}{100}n$ of vertices $j \in V$ and corresponding vectors we have $v_o^* v_j^* > 0.0004$, then $z^{2*} \geq \frac{n}{2} + \frac{3n}{2000000}$. Therefore, based on the Assumption (3. b) for all feasible solutions $V = V_{1G} \cup V_{-1G}$ we have $\frac{|V_{1G}|}{z^{2*}} \leq 1.999999$. Otherwise, go to Step 4.

Step 4. Introduce two normalized random vectors u and w, and produce H_u and H_w. If $|H_u| \geq \frac{n}{1000000}$ or $|H_w| \geq \frac{n}{1000000}$ then produce the suitable solution $V_{1G} \cup V_{-1G}$, correspondingly. Therefore, based on the Corollary (3) we have $\frac{|V_{1G}|}{z^{2*}} \leq 1.999999$.

Corollary 4. Based on the proposed 1.999999-approximation algorithm for the vertex cover problem and by assuming the unique games conjecture, then $P = NP$.

4. Conclusions

One of the open problems about the vertex cover problem is the possibility of introducing an approximation algorithm within any constant factor better than 2. Here, we proposed a new algorithm to introduce a 1.999999-approximation algorithm for the vertex cover problem on arbitrary graphs, and this may lead to the conclusion that $P = NP$.

References

1. Dinur, I., & Safra, S. (2005). On the hardness of approximating minimum vertex cover. *Annals of Mathematics*, 162, 439-485.

2. Khot, S., (2002, May). On the power of unique 2-Prover 1-Round games. Proceeding of 34th ACM Symposium on Theory of Computing, STOC.

3. Khot, S., & Regev, O. (2008). Vertex cover might be hard to approximate to within 2-Ɛ. *Journal of Computer and System Sciences*, 74, 335-349.