New insight into introducing a $(2 - \mathcal{E})$-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. 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 introducing a modified graph and corresponding model along with satisfying the proposed assumptions, we propose new insight into solving this open problem and we introduce an approximation algorithm with a performance ratio of 1.99997 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 $(2 - \mathcal{E})$-approximation ratio on special graphs, and then, we show that on arbitrary graphs a $(2 - \mathcal{E})$-approximation ratio can be obtained by solving a new semidefinite programming problem (SDP). 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 solve a new SDP model 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 graph G) and objective value $|V_{1G}|$.

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

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

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

And by considering the triangle inequalities, we can introduce the following integer linear programming (ILP) model for the minimum vertex cover problem:

$$(1) \qquad \min_{s.t.} \quad z1=\sum_{1\le j\le n}\frac{1+x_j}{2}$$

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

Here, triangle inequalities are as cutting plane inequalities, and 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:

$$(2) \qquad \min_{s.t.} \quad z2 = \sum_{i \in V} \frac{1 + v_o v_i}{2}$$

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

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

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

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

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

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

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

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

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

Assumption 1. From now on, we assume that $\frac{n}{2} \leq z^* < \frac{n}{2} + \frac{n}{100000}$.

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 the approximation ratio $\frac{|V_{1G}|}{Z^*} \leq \frac{2k}{k+1}$.

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

Therefore, for bounded values of k, we have some approximation ratios smaller than 2. But, if $k \to \infty$ then $\frac{|V_{1G}|}{Z^*} \to 2$, and we don't have a constant approximation ratio better than 2.

Corollary 1. If $|V_{1G}| < \frac{k}{k+1}n$ then $|V_{1G}| < k|V_{-1G}|$ and $\frac{|V_{1G}|}{Z^*} < \frac{2k}{k+1}$.

Assumption 2. We don't have a suitable feasible solution $\mathrm{V} = \mathrm{V}_1 \cup \mathrm{V}_{-1}$ for which $|\mathrm{V}_1| < 0.999n$; i.e. We assume that we have $V = V_1 \cup V_{-1}$ as an arbitrary feasible solution for which $|V_1| \geq 0.999n$.

Up to now, we could introduce $(2-\varepsilon)$-approximation ratio on special graphs with suitable characteristics. In section 3, we are going to introduce such a ratio on arbitrary graphs.

3. A $(2 - \varepsilon)$-Approximation algorithm for 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 1.99997-approximation ratio on arbitrary graphs. To do this, let $\varepsilon = 0.001$ and suppose that we have a graph $G(V,E)$ with a feasible solution $V_1 \cup V_{-1}$, where $|V_1| \geq 0.999n$ and $\frac{n}{2} \leq z^* < \frac{n}{2} + \frac{n}{100000}$.

Assumption 3. By solving the SDP relaxation (2), for more than $0.9n$ of vertices $j \in V$ and corresponding vectors we have $-\varepsilon \leq v_0^* v_j^* \leq +\varepsilon$. Because for less than $0.001n$ of vertices $j \in V$ and corresponding vectors we have $v_o^* v_j^* < 0$ (we can't produce a feasible solution with $|V_1| < 0.999n$), and $z^* < \frac{n}{2} + \frac{n}{100000}$. Let $G_\varepsilon = \{j \in V | -\varepsilon \leq v_o^* v_j^* \leq +\varepsilon\}$, where $|G_\varepsilon| \geq 0.9n$.

Theorem 3. For a vertex $k \in V$ and the corresponding set $H_k = \{j \in G_\varepsilon;\ |v_k^* v_j^*| > \varepsilon\}$, H_k is a bipartite graph.

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

$$S = \{j \in H_k |\ v_k^* v_j^* < -\varepsilon\} \quad \text{and} \quad T = \{j \in H_k |\ v_k^* v_j^* > +\varepsilon\}$$

Then, it is sufficient to show that the sets S and T are null subgraphs. Based on the first constraint of the SDP model (2), we have $v_i^* v_j^* \leq -1 + 2\varepsilon \quad ij \in E(G),\ i,j \in H_k$. Therefore, if $i,j \in S$ then the second constraint of the SDP model (2) is violated and if $i,j \in T$ then the third constraint is violated ■

Corollary 2. If $\exists k \in V: |H_k| \geq \frac{n}{1000}$ then we have a feasible solution $V_{1G} \cup V_{-1G}$, correspondingly, where $|V_{-1G}| = max\{|S|,|T|\} \geq \frac{n}{2000}$. Hence, $|V_{1G}| \leq 1999|V_{-1G}|$ and $\frac{|V_{1G}|}{z^*} \leq \frac{2\times1999}{1999+1} = 1.999 < 2$.

Assumption 4. $\forall k \in V: |H_k| < \frac{n}{1000}$ and we can't produce a suitable feasible solution. In this case, for each vector v_k^*, it is almost perpendicular to most of the vectors v_j^*. Moreover, we can show that each vertex $k \in V$ has a long-distance (on G_ε) to most of the vertices of G_ε and this is our reason to introduce the following graph.

Definition 1. For each pair of vertices i and j of graph $G = (V,E)$, add two new vertices i_j and j_i and a path with distance 3 through these vertices to produce the corresponding graph $H_G = (V \cup V', E \cup E')$, where $V' = \{i_j, j_i | i,j \in V\}$ and $E' = \{(i,i_j),(i_j,j_i),(j_i,j) | i,j \in V\}$, $|V'| = 2\binom{|V|}{2}$ and $|E'| = 3\binom{|V|}{2}$.

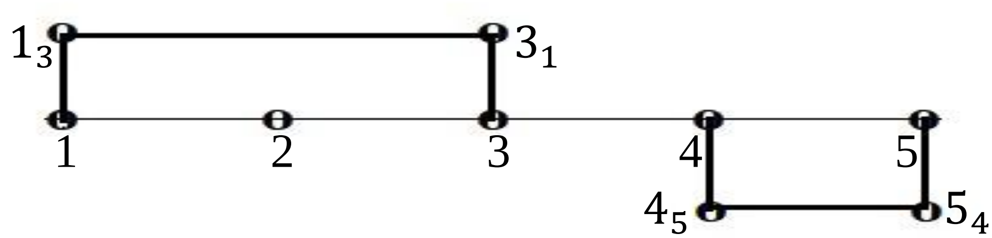

Figure 1. Addition of the new vertices to construct G'.

Now, we add $|V'|$ more 'shadow' points to a new SDP model so that for every unit vector $v_{i_j}\ \ i,j \in V$ we add unit vector $v_{i'_j}$ which has an angle of $\frac{\pi}{3}$ degree with v_{i_j}; i.e. $v_{i_j} v_{i'_j} = \frac{1}{2}$.

Corollary 3. a) $v_{i'_j}v_{j'_i} = -1$ if and only if $v_{i_j}v_{j_i} \in \{-1, +\frac{1}{2}\}$,

b) $v_{i'_j}v_{j'_i} = -\frac{1}{2}$ if and only if $v_{i_j}v_{j_i} \in \{-\frac{1}{2}, +1\}$.

Then, we introduce the following SDP model, which is almost similar to the SDP model (2). Here, the objective function of the model (2) is embedded into the constraints of the model (3), and its objective function is introduced so that for most of the pairs of vertices $i, j \in V$ we have $v_{i_j}v_{j_i} = v_{i'_j}v_{j'_i} = -1$.

$$
\begin{aligned}
(3) \qquad \min_{s.t.} \quad & z3 = \sum_{i,j\in V} \frac{v_{i_j}v_{j_i} - 1}{2} + \sum_{i,j\in V} \frac{v_{i'_j}v_{j'_i} - 1}{2} \\
& +v_o v_i + v_o v_j - v_i v_j = 1 \qquad ij \in E \cup E' \\
& +v_i v_j + v_i v_k + v_j v_k \geq -1 \qquad i,j,k \in V \cup V' \cup \{o\} \\
& +v_i v_j - v_i v_k - v_j v_k \geq -1 \qquad i,j,k \in V \cup V' \cup \{o\} \\
& -v_i v_j + v_i v_k - v_j v_k \geq -1 \qquad i,j,k \in V \cup V' \cup \{o\} \\
& -v_i v_j - v_i v_k + v_j v_k \geq -1 \qquad i,j,k \in V \cup V' \cup \{o\} \\
& \sum_{i\in V} \frac{1 + v_o v_i}{2} \leq \hat{z} = \left\lceil \frac{|V|}{2} + \frac{|V|}{100000} \right\rceil \\
& v_{i_j}v_{i'_j} = v_{j_i}v_{j'_i} = \frac{1}{2} \qquad i,j \in V \\
& v_i v_i = v_{i'_j}v_{i'_j} = 1 \qquad i,j \in V \cup V' \cup \{o\} \\
& v_i v_j \in \{-1, +1\} \qquad i,j \in V \cup V' \cup \{o\}
\end{aligned}
$$

Corollary 4. If we have a feasible vertex cover on graph G with $|V_{1G}| = \hat{z} \in \mathbb{N}$, then in the optimal solution of SDP (3) relaxation we have: $z3^* \leq -2\left(\binom{|V|}{2} - \binom{|V| - \hat{z}}{2}\right) - \frac{1}{2}\underbrace{\left(\binom{|V| - \hat{z}}{2}\right)}_{i,j\in V_{-1G}}$.

Theorem 4. If we have a feasible vertex cover on graph G with $|V_{1G}| \leq \hat{z} \in \mathbb{N}$, then in the optimal solution of SDP (3) relaxation we have: $v_{i_j}v_{j_i} \in \{-1, +1\}$ and $v_{i'_j}v_{j'_i} \in \left\{-1, -\frac{1}{2}\right\}$ $i, j \in V$.

Proof. Based on the induction on n. It is true for $n = 2$. Suppose that it is true for n and we should prove it for a graph $G = (V, E)$, where $|V| = n + 1$. Let $\mathrm{k} \in \mathrm{V}$ as a vertex which is in a vertex cover solution of G with objective value $|V_1| = \hat{z}$. By removing the vertex k and all vertices in V' which had been introduced based on vertex k, we have a new model on $G' = G_{V-\{k\}}$ with $\hat{z} - 1$ as the value for the right-hand side of the sixth constraint, where $z3^{*new} \leq z3^{*old} + 2n$.

Therefore, we have an optimal solution of SDP (3) relaxation on G' for which $v_{i_j}v_{j_i} \in \{-1, +1\}$ and $v_{i'_j}v_{j'_i} \in \left\{-1, -\frac{1}{2}\right\}$ $i, j \in V$-$\{\mathrm{k}\}$. Then, by setting $v_k = -v_{k_j} = v_{j_k} = +v_o^*$ $j \in V - \{k\}$ and introducing the vectors $v_{k'_j} = -v_{j'_k}$ $j \in V - \{K\}$ such that $v_{k'_j} v_o^* = \frac{1}{2}$, we have a feasible solution on graph G with objective value $z3^{*new} - 2n \leq z3^{*old}$ that is an optimal solution of SDP (3) relaxation on graph G for which $v_{i_j} v_{j_i} \in \{-1, +1\}$ and $v_{i'_j} v_{j'_i} \in \left\{-1, -\frac{1}{2}\right\}$ $i, j \in V$ ■

This means that in the optimal solution of SDP (3) relaxation, for less than $\binom{|V| - \hat{z}}{2} = \binom{n - \hat{z}}{2}$ pairs of vertices $i, j \in V$ we have $v_{i_j} v_{j_i} = +1$. Now, suppose that after solving the SDP (3) relaxation, for more than $0.9n$ of vertices $j \in V$ and corresponding vectors we have $-\varepsilon \leq v_o^* v_j^* \leq +\varepsilon$; i.e. $|G_\varepsilon| \geq 0.9n$ (Assumption 3). Therefore, for more than $\binom{0.9n - (n - \hat{z})}{2} = \binom{-0.1n + \hat{z}}{2}$ pairs of vertices $i, j \in G_\varepsilon$ we have $v_{i_j} v_{j_i} = -1$. Hence, if $\hat{z} = \left[\frac{|V|}{2} + \frac{|V|}{100000}\right]$ then for more than $\binom{0.4n}{2} \cong .08\mathrm{n}^2$ pairs of these vertices we have $v_{i_j} v_{j_i} = -1$ which concludes that $v_i v_j \cong -1 < -\varepsilon$ for all of these paired vertices.

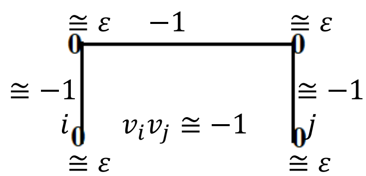

Figure 2. Each of paired vectors v_i , v_{i_j} and v_{i_j} , v_{j_i} and v_{j_i} , v_j and consequently v_i , v_j have almost an angle of π degree with each other.

Theorem 5. If we have a feasible vertex cover on graph G with $|V_{1G}| \leq \hat{z} \in \mathbb{N}$, then in the optimal solution of SDP (3) relaxation $\exists k \in V$: $|H_k| \geq \frac{n}{1000}$.

Proof. If $\forall k \in V$: $|H_k| < \frac{n}{1000}$ then for at most $\frac{\mathrm{n}}{2} \times \frac{n}{1000} = 0.0005\mathrm{n}^2$ pairs of vertices $i, j \in V$ we have $|v_i v_j| > \varepsilon$ which is a contradiction ($0.0005\mathrm{n}^2 \ll .08\mathrm{n}^2$) ■

Hence, if we have a feasible vertex cover on graph G with $|V_{1G}| \leq \hat{z} \in \mathbb{N}$, then $\exists k \in V$: $|H_k| \geq \frac{n}{1000}$ and based on the optimal solution of SDP (3) relaxation we can produce a suitable feasible solution $V_{1G} \cup V_{-1G}$, correspondingly, where $|V_{-1G}| \geq \frac{n}{2000}$. Hence, $\frac{|V_{1G}|}{z^*} \leq \frac{2 \times 1999}{1999+1} = 1.999 < 2$.

Now, we can introduce an algorithm to produce an approximation ratio $\rho \leq 1.99997$.

Zohrehbandian Algorithm (To produce a vertex cover solution with a factor $\rho \leq 1.99997$)

Step 1. Solve the SDP (3) and produce $\mathrm{H_k}$'s.

Step 2. If $\exists k \in V$: $|H_k| \geq \frac{n}{1000}$ then produce the suitable solution $\mathrm{V_{1G} \cup V_{-1G}}$, correspondingly, where $|\mathrm{V_{-1G}}| = \max\{|\mathrm{S}|, |\mathrm{T}|\}$. Hence, $|\mathrm{V_{1G}}| \leq 1999|\mathrm{V_{-1G}}|$ and $\frac{|\mathrm{V_{1G}}|}{\mathrm{z}^*} \leq 1.999$. Otherwise, the optimal vertex cover is greater than $\frac{|\mathrm{V}|}{2} + \frac{|\mathrm{V}|}{100000}$ and for all feasible solutions, we have $\frac{|\mathrm{V_{1G}}|}{\mathrm{Z}^*} \leq \frac{\mathrm{n}}{\mathrm{Z}^*} \leq \frac{200000}{100002} < 1.99997$.

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 1.99997-approximation algorithm for the vertex cover problem on arbitrary graphs.

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.