On Some Approximation Algorithms for Dense Vertex Cover Problem

Anton V. Eremeev, Omsk Branch of Sobolev Institute of Mathematics, RAS

Summary: In this work we study the performance of some algorithms approximating the vertex cover problem (VCP) with respect to density constraints. We consider a modification of the algorithm of Karpinski and Zelikovsky [9] for the weighted VCP and prove a performance guarantee for this algorithm in terms of a new density parameter. Also we investigate the hardness of approximation for the VCP on everywhere ε-dense graph and show that it is NP-hard to approximate within a factor less than $\frac{7+\varepsilon}{6+2\varepsilon}$.

1. Introduction

Let $G = (V, E)$ be a graph with the set of vertices V and the edge set E. A subset $C \subseteq V$ is called a vertex cover of G if every edge has at least one endpoint in C. The VCP is: given a graph $G = (V, E)$ and weight $w_v \geq 0$ for each $v \in V$, find a vertex cover C^* with minimum total weight $w(C^*)$. The unweighted VCP is a case when $w_v = 1$ for all $v \in V$. Here we denote $w(U) = \sum_{u \in U} w_u$ for any $U \subseteq V$.

In [7] Håstad shows that for VCP there is no polynomial δ-approximation algorithm if $\delta < 7/6$, i.e. no algorithm finds a cover of weight less than 7/6 of optimum in polynomial time, unless $P = NP$. Bar-Yehuda and Even proposed a linear time 2-approximation algorithm for this problem in [2]. No polynomial algorithms with better constant ratio approximation bound are known for the general case of VCP, although various polynomial algorithms with stronger performance guarantees are found for different restricted cases (see e.g. [8]).

Let us start with a definition of *everywhere ε-dense* graph in which every vertex is required to have at least $\varepsilon|V|$ neighbors for some fixed parameter $\varepsilon > 0$. Arora et al [1] prove that on these graphs a number of NP-hard problems, such as the maximum cut problem, admit the polynomial time approximation scheme, i.e. the approximation ratio δ may be achieved for them in polynomial time for any $\delta > 1$. These results do not apply to the VCP however. In fact, Clemetni and Trevisan [5] show that even in case of unweighted VCP on the everywhere ε-dense graph for any $\varepsilon \in (0, 1)$ there exists such $\Delta(\varepsilon) > 1$ that obtaining a cover with approximation ratio less than $\Delta(\varepsilon)$ is NP-hard. An example of the function $\Delta(\varepsilon)$ will be given explicitly further in this paper.

In [9] Karpinski and Zelikovsky designed an approximation algorithm for the unweighted case of VCP with a guaranteed performance ratio $\frac{2}{1+\varepsilon}$. We shall refer to this algorithm as DVC. In this paper we will study how the concept of density could be extended to the weighted case permitting a similar approximation guarantee.

2. Performance Guarantees in Terms of Density Parameters

We consider the following modification of the DVC algorithm. Let $N(v)$ be the set of vertices adjacent to v, and let $G(V')$ denote a graph induced by a vertex set $V' \subseteq V$. The algorithm implements a reduction rule based on the simple observation that if a vertex v does not belong to a cover C then

$N(v) \subseteq C$. Once the neighborhood is fixed, the approximate solution C_v to the remaining VCP on $G(V \backslash (N(v) \cup \{v\}))$ may be found by some α-approximation algorithm A_α.

Algorithm DVC$_\alpha$

For all $v \in V$ do

$V' \leftarrow V \backslash (N(v) \cup \{v\})$.

Find a vertex cover C_v for $G(V')$ using the algorithm A_α.

$C(v) \leftarrow N(v) \cup C_v$.

Return $C = argmin_{v \in V} w(C(v))$.

The DVC algorithm [9] may be considered as DVC$_2$, where the 2-approximation algorithm [2] is used as A_α algorithm.

In case of weighted VCP on the everywhere ε-dense graph the DVC$_\alpha$ scheme does not yield a substantial improvement to the performance guarantee of A_α. Indeed suppose that $L_1, L_2, \ldots$ are the copies of the graph for which the approximation factor α is attained by A_α. Denote the size of the vertex set of this graph by $|V_L|$. Let us consider an n-vertex graph G, which consists of $k(n) = \left\lfloor \frac{n(1-\varepsilon)}{|V_L|} \right\rfloor$ subgraphs $L_1, L_2, \ldots, L_{k(n)}$ and a clique on $n - k(n) \cdot |V_L| \geq \varepsilon n$ vertices. Join each vertex of $L_1, L_2, \ldots, L_{k(n)}$ with all vertices of the clique and assume that the weight of the clique vertices equals 0. The graph G is everywhere ε-dense, but DVC$_\alpha$ is able to improve the approximation ratio α for not more than one of subgraphs L_i. Hence, $\frac{w(C)}{w(C^*)} \geq \frac{\alpha(k(n)-1)}{k(n)} \underset{n \to \infty}{\to} \alpha$.

To formulate the conditions that guarantee the approximation factor for DVC$_\alpha$ which is better than α, we will need a new definition of density. We say that a graph G is a *ρ-dense weight graph* if $w(N(v)) \geq \rho w(V)$ for any vertex $v \in V$. The following theorem extends the result from [9] to the weighted case.

Theorem 1 *The weight of the vertex cover returned by DVC$_\alpha$ algorithm for a ρ-dense weight graph is at most $\frac{\alpha}{1+\rho(\alpha-1)}$ times the weight of the optimal cover.*

Proof: Let $v \in V \backslash C^*$. Then $N(v) \subseteq C^*$ and the vertices of $N(v)$ cover all edges between $N(v)$ and $V' = V \backslash (N(v) \cup \{v\})$. So the rest of the vertices in C^* give the optimal cover for $G(V')$.

Denote $C_v^* = C^* \backslash N(v)$ and let C_v be the cover of $G(V')$ returned by the algorithm A_α. Then $w(C_v) \leq \alpha w(C_v^*)$ and therefore,

$$\frac{w(C)}{w(C^*)} \leq \frac{w(N(v)) + \alpha w(C_v^*)}{w(N(v)) + w(C_v^*)} \leq \frac{\rho w(V) + \alpha w(C_v^*)}{\rho w(V) + w(C_v^*)} = \alpha - \frac{\alpha - 1}{1 + w(C_v^*)/(\rho w(V))}.$$

If $\alpha w(C_v^*) \leq (1-\rho) w(V)$ then

$$\frac{w(C)}{w(C^*)} \leq \alpha - \frac{\alpha - 1}{1 + \frac{(1-\rho)w(V)}{\alpha \rho w(V)}} = \frac{\alpha}{1 + \rho(\alpha - 1)}.$$

If $\alpha w(C_v^*) > (1-\rho) w(V)$ then the inequality $w(C) \leq w(V)$ also yields the desired bound:

$$\frac{w(C)}{w(C^*)} \leq \frac{w(V)}{w(N(v)) + w(C_v^*)} < \frac{w(V)}{\rho w(V) + \frac{(1-\rho)w(V)}{\alpha}} = \frac{\alpha}{1 + \rho(\alpha - 1)}. \qquad \square$$

Note that if $w_v > 0$ for all $v \in V$ then the proved inequality is strict. Indeed, the equality would imply that on one hand, $\alpha w(C_v^*) = (1-\rho)w(V)$, on the other, $w(N(v)) = \rho w(V)$, and $\alpha w(C_v^*) = w(C_v)$. Thus, $\alpha w(C_v^*) \leq w(V') = w(V) - \rho w(V) - w_v = \alpha w(C_v^*) - w_v$, which is impossible if $w_v > 0$.

The theorem shows that DVC_2 is a $\frac{2}{1+\rho}$-approximation algorithm for the VCP on ρ-dense weight graph. Clearly, in case if all weights equal 1, the parameters ρ and ε coincide.

Corollary *If d is such that $w_u \leq dw_v$ holds for any vertices u and v of G, and graph G is everywhere ε-dense then the approximation ratio of DVC_2 does not exceed $2 - \frac{2}{2+(\varepsilon^{-1}-1)d}$.*

Proof: Note that

$$\frac{w(V)}{w(N(v))} = 1 + \frac{w(V) - w(N(v))}{w(N(v))} \leq 1 + \frac{|V \backslash N(v)| \max\limits_{u \in V} w_u}{|N(v)| \min\limits_{u \in V} w_u} \leq 1 + (\varepsilon^{-1} - 1)d.$$

Application of the theorem gives the required approximation bound. □

Let n denote the number of vertices in graph G. The $\left(2 - \frac{\log \log n}{2 \log n}\right)$-approximation algorithm suggested by Bar-Yehuda and Even [3] may be used as A_α in DVC_α as well. This gives a polynomial-time algorithm for the ρ-dense weight graph with approximation ratio $\frac{4 \log n - \log \log n}{2(\rho+1)\log n - \rho \log \log n}$, which is less than $\frac{2}{1+\rho}$.

In order to estimate the gap between the approximation factor achieved and the approximation which is NP-hard for the everywhere ε-dense graph, we have combined the above-mentioned results of Clementi and Trevisan [5] and Håstad [7].

3. Hardness of Approximation

Here we will use the following standard notation (for the detailed definitions see e.g. [4, 5]). Let the *verifier* be a probabilistic polynomial-time Turing Machine using the random string tape R and an oracle *proof string* tape π. The verifier reads the input x of size $\mathbf{n}$ and the random string of at most $O(r(\mathbf{n}))$ random bits. The proof π is examined in certain locations depending on x and R. Finally the input is either accepted or rejected by the verifier. The *accepting probability* of input x is the maximum over all proofs π of probability that x is accepted.

The verifier is restricted by a number of parameters. Let f be the *free bit complexity* of the verifier if for any input x and any random string R there are at most 2^f possible accepting configurations (x, R, a), where a is the sequence of bits received from the proof (given x and R the set of all such sequences should be polynomially obtainable). The verifier for language L has a *soundness* s if for any $x \notin L$ its accepting probability is not more than s. The verifier has a *completeness* c if for any $x \in L$ the accepting probability is not less than c.

A language L belongs to the class $FPCP_{c,s}[r(\mathbf{n}), f]$ where $0 < s < c \leq 1, f > 0$, if there exists a verifier with a free bit complexity f, soundness s, completeness c, and $O(r(\mathbf{n}))$ random bits for L. In this paper we assume that f is integer. Using the approach proposed in [5] we obtain the following.

Lemma 1 *Suppose that $NP \subseteq FPCP_{c,s}[\log(\mathbf{n}), f]$. Then it is NP-hard to p-approximate the VCP on everywhere ε-dense graph, if p is such that*

$$p + \frac{\varepsilon(p-1)}{(1-\varepsilon)(1-2^{-f})} < 1 + \frac{c-s}{2^f - c}. \tag{1}$$

Proof: Let us consider the VCP on $G_\phi = (V_\phi, E_\phi)$, which is an FGLSS graph (see e.g. [5, 6]) constructed for the verifier with the given parameters $c, s, f, r(\mathbf{n}) = \log(\mathbf{n})$ and an instance ϕ of the satisfiability problem. In this graph each vertex corresponds to one of the accepting configurations (x, R, a). Adding dummy configurations if it is necessary we assume that $|V_\phi| = \mathbf{r} \cdot 2^f$, where $\mathbf{r} = 2^{O(\log(\mathbf{n}))}$ is the total number of possible random sequences that can be accessed by the verifier. A pair of vertices in G_ϕ is adjacent if no such proof π exists that complies with the accepting configurations of both vertices, i.e. the examined parts of proofs matching these configurations should always contain some different bits.

For arbitrary $\varepsilon \in (0,1)$ we join each vertex of G_ϕ to all vertices of an additional clique which has a vertex set Q of size $\left\lfloor \frac{\varepsilon |V_\phi|}{1-\varepsilon} \right\rfloor$. The obtained graph $H_\phi(\varepsilon)$ is everywhere ε-dense (note that the vertices corresponding to the same random string R are adjacent and each $v \in V_\phi$ has a neighbor in G_ϕ).

Let C' be a p-approximate solution to the VCP on $H_\phi(\varepsilon)$ and let C^*_G and C^*_H denote the optimal covers for G_ϕ and $H_\phi(\varepsilon)$ respectively. Here we assume that $C' \cap Q = C^*_H \cap Q = Q$ (otherwise a cover may be easily substituted by a new cover of the same cardinality matching this condition). A vertex set $C = C' \backslash Q$ is a cover for G_ϕ, and $|C| = |C'| - |Q|$. Since $|C^*_G| = |C^*_H| - |Q|$, so $|C'| \le p(|C^*_G| + |Q|)$, and $|C| \le p|C^*_G| + p|Q| - |Q|$. It is not difficult to see that the size of the largest independent set I^*_G in G_ϕ does not exceed $\mathbf{r} = \frac{|V_\phi|}{2^f}$. Therefore, since $|C^*_G| = |V_\phi| - |I^*_G|$, we have $|C^*_G| \ge |V_\phi| - |V_\phi| 2^{-f}$, and

$$\frac{|C|}{|C^*_G|} \le p + \left\lfloor \frac{\varepsilon |C^*_G|}{(1-\varepsilon)(1-2^{-f})} \right\rfloor \frac{p-1}{|C^*_G|}. \tag{2}$$

The proof of VCP non-approximability in [4, 5] implies that if $NP \subseteq FPCP_{c,s}[log(\mathbf{n}), f]$, then $\left(1 + \frac{c-s}{2^f - c} - \delta\right)$-approximation for the VCP on G_ϕ is NP-hard for any $\delta > 0$. In case if (1) holds, any p-approximate solution C' to the VCP on everywhere ε-dence graph $H_\phi(\varepsilon)$ provides a cover C for G_ϕ which is NP-hard to obtain due to (2). □

Theorem 2 *It is NP-hard to find p-approximate solution to the VCP on everywhere ε-dense graph if $p < \Delta(\varepsilon) = \frac{7+\varepsilon}{6+2\varepsilon}$.*

Proof: By [7] $NP \subseteq FPCP_{1-\mu, 0.5+\nu}[\log(\mathbf{n}), 2]$ if $\mu > 0, \nu > 0$. Thus, $1 + \frac{c-s}{2^f - c} = \frac{7}{6} - \delta$, where $\delta > 0$ can be made arbitrarily close to zero. Application of the lemma yields the required bound. □

The theorem implies that in general attaining the approximation factor less than $\frac{7+\rho}{6+2\rho}$ for the ρ-dense weight graph is NP-hard.

4. Conclusions

We have investigated the potentialities of an approximation scheme based on a simple problem reduction rule in combination with an approximation algorithm. The analysis rendered a new density parameter defining the approximation ratio in case of the weighted VCP. The considered reduction rule should be particularly efficient in heuristic algorithms. A forthcoming paper will describe a version of genetic algorithm where the mutation and crossover operators are based on this rule.

The analysis of hardness of approximation leaves a gap between the approximation ratio of DVC_2 and the factor which is provably NP-hard to obtain. However the gap decreases to zero as the density parameter grows.

Acknowledgment

The author would like to thank M.I. Sviridenko for the helpful comments on the paper.

The research was supported in part by INTAS Grants 96-0820, 99-3949 and Russian Foundation for Basic Research Grant 97-01-00771.

Corresponding Author: Anton Eremeev, Omsk Branch of Sobolev Institute of Mathematics,
13 Pevtsov st. 644099, Omsk, Russia.
E-mail: eremeev@iitam.omsk.net.ru

References

[1] Arora, S.; Karger, D.; Karpinski, M. (1995): Polynomial Time Approximation Schemes for Dense Instances of *NP*-Hard Problems. In: Proceedings of 27th ACM Symposium on Theory of Computing, pp. 284-293.

[2] Bar-Yehuda, R.; Even, S. (1981): A Linear Time Approximation Algorithm for the Weighted Vertex Cover Problem. Journal of Algorithms, Vol. 2. pp. 198-203.

[3] Bar-Yehuda, R.; Even, S. (1985): A Local-Ratio Theorem for Approximating the Weighted Vertex Cover Problem. Annals of Discrete Mathematics, Vol. 25. pp. 27-45.

[4] Bellare, M.; Goldreich, O.; Sudan, M. (1995): Free Bits and Non-Approximability - Towards Tight Results. In: Proceedings of 36th Annual IEEE Symposium on Foundations of Computer Sciense. pp. 422-431. Full version available in ECCC Technical Report TR 95-24.

[5] Clementi, A.E.F.; Trevisan, L. (1996): Improved Non-Approximability Results for Vertex Cover with Density Constraints. In: Proceedings of 2nd Computing and Combinatorics Conference. COCOON'96, Berlin, Springer Verlag, pp. 333-342.

[6] Feige, U.; Goldwasser, S.; Lovász, L.; Safra, S.; Szegedy, M. (1991): Approximating Clique is Almost NP-complete. In: Proceedings of the 32nd IEEE Symposium on Foundations of Computer Science. pp 2-12.

[7] Håstad, J. (1997): Some Optimal Inapproximability Results. Electronic Colloquium on Computational Complexity, ECCC Technical Report TR 97-037.

[8] Hochbaum, D.S. (1995): Approximating Covering and Packing Problems: Set Cover, Vertex Cover, Independent Set, and Related Problems. In: Approximation Algorithms for NP-Hard Problems. Ed. by S.D.Hochbaum, PWS Publishing Company, pp. 94-143.

[9] Karpinski, M.; Zelikovsky, A. (1996): Approximating Dense Cases of Covering Problems. DIMACS Technical Report TR 96-59. International Computer Science Institute, Berkeley.