Theoretical Computer Science 555 (2014) 2–8

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Approximation algorithms for the partition vertex cover problem

Suman K. Bera [a,*,1], Shalmoli Gupta [b,1], Amit Kumar [c], Sambuddha Roy [d]

[a] Dartmouth College, United States
[b] University of Illinois at Urbana-Champaign, United States
[c] Indian Institute of Technology Delhi, India
[d] IBM India Research Lab, New Delhi, India

ARTICLE INFO

Article history:
Received 20 April 2013
Received in revised form 18 March 2014
Accepted 2 April 2014
Available online 8 April 2014

Keywords:
Approximation algorithm
Vertex cover
Knapsack cover inequality
LP rounding

ABSTRACT

We consider a natural generalization of the `Partial Vertex Cover` problem. Here an instance consists of a graph $G=(V,E)$, a cost function $c: V \to \mathbb{Z}^+$, a partition $P_1,\dots,P_r$ of the edge set E, and a parameter k_i for each partition P_i. The objective is to find a minimum cost set of vertices which cover at least k_i edges from the partition P_i. We call this the `Partition-VC` problem. In this paper, we give matching upper and lower bound on the approximability of this problem. Our algorithm is based on a novel LP relaxation for this problem. This LP relaxation is obtained by adding knapsack cover inequalities to a natural LP relaxation of the problem. We show that this LP has integrality gap of $O(\log r)$, where r is the number of sets in the partition of the edge set. We also extend our result to more general settings. For example we consider a problem where additionally edges have profits, and we would like to pick a minimum cost set of vertices which cover edges of total profit at least Π_i for each partition P_i. We call this the `Knapsack Partition Vertex Cover` problem. We give an $O(\log r)$ approximation algorithm for this problem as well.

1. Introduction

The `Vertex Cover` problem is one of the most fundamental NP-hard problems and has been widely studied in the context of approximation algorithms [1,2]. In this problem, we are given an undirected graph $G=(V,E)$ and a cost function $c: V \to \mathbb{Z}^+$. The goal is to find a minimum cost set of vertices which *cover* all the edges in E: a set of vertices S covers an edge e if S contains at least one of the end-points of e. Several 2-approximation algorithms are known for this problem [3,4]. The `Partial Vertex Cover` problem is a generalization of the `Vertex Cover` problem, where we are also given a parameter k. The goal is to find a minimum cost set of vertices which cover at least k edges. This problem was proposed by Bshouty and Burroughs [5], and they gave a 2-approximation for this problem using LP-rounding. Since then, many different techniques have been shown to give 2-approximation algorithm for this problem [6–8].

In this paper, we consider a natural generalization of the `Partial Vertex Cover` problem. Here an instance consists of a graph $G=(V,E)$, a cost function $c: V \to \mathbb{Z}^+$, a partition $P_1,\dots,P_r$ of the edge set E, and a parameter k_i for each

* Corresponding author.
E-mail addresses: sumankalyanbera@gmail.com (S.K. Bera), shalmoli@gmail.com (S. Gupta), amitk@cse.iitd.ernet.in (A. Kumar), sambuddha@in.ibm.com (S. Roy).

[1] Most of this work was done while the author was a student at Indian Institute of Technology Delhi.

http://dx.doi.org/10.1016/j.tcs.2014.04.006
0304-3975/

partition P_i. The goal is to find a minimum cost set of vertices which cover at least k_i edges from the partition P_i. We call this the Partition-VC problem. In this paper, we give matching upper and lower bound on the approximability of this problem. We give an $O(\log r)$-approximation algorithm for the Partition-VC problem, and show that unless P = NP, we cannot do better. Recall that r denotes the number of sets in the partition of the edge set E. Note that for constant number of partitions this gives a constant approximation. Our results also extend to the slightly more general problem where edges additionally have profits, and we would like to pick a minimum cost set of vertices which cover edges of total profit at least Π_i for each partition P_i. We call this the Knapsack Partition Vertex Cover problem.

Our techniques: The hardness result for the Partition-VC problem follows by an approximation preserving reduction from the Set Cover problem. The approximation algorithm uses a novel LP relaxation – this is the main contribution of the paper, and we expect this idea to have more applications. The natural LP relaxation for even the Partial Vertex Cover turns out to be unbounded. Indeed, consider the following example: the graph is a star – there is a vertex v of degree D and all its neighbors are leaves. All vertex costs are 1, and the parameter $k = 1$. Clearly, any optimal solution must cost at least 1 unit. But a fractional solution will pick the vertex v to an extent of $\frac{1}{D}$, and hence will cover all the D edges fractionally to an extent of $\frac{1}{D}$. So, the fractional solution pays only $\frac{1}{D}$. One way of getting around this problem is to augment the LP with more information. Here, we can guess the most expensive vertex an optimal solution will pick, and can remove all vertices with cost more than the cost of this vertex. Further, the cost of this vertex is also a lower bound on the optimal value. This idea was used by [8] to give a 2-approximation for the Partial Vertex Cover problem. However, applying such an idea to the Partition-VC problem turns out to be non-trivial. We cannot guess the most expensive vertex in each of the partitions – this will take time exponential in r. Our approach is to strengthen the natural LP relaxation such that no guesswork is required. We show how to do this using knapsack cover inequalities [9]. Armed with this stronger relaxation, we show that one can carefully use randomized rounding based techniques to get the approximation algorithm.

Related work: There has been recent work on partial covering versions of several covering problems. For the set cover problem, the partial covering version namely the partial set cover problem was first studied by Kearns [10], who proved that the approximation ratio of the greedy algorithm is at most $2H(n) + 3$, where n is the size of the ground set and $H(n)$ is the nth harmonic number. Later Slavík [11] showed that it is actually bounded by $H(k)$ where k is the number of elements to be covered. This natural greedy approach when extended for the Partition-VC problem gives only an $H(|V|)$-approximation, which is much worse than the lower bound of $O(\log r)$ that we have proved for this problem.

The partial vertex cover problem has also been widely studied in the literature. Bshouty and Burroughs [5] were the first to give a polynomial time 2-approximation algorithm for it. Subsequently, several other algorithms based on Lagrangian Relaxation, local-ratio, primal-dual techniques with the same approximation guarantee were proposed [6–8,12]. Mestre's [12] primal-dual technique can also be used to get a 2-approximation for a more general version of the problem, the partial capacitated vertex cover problem. Bar-Yehuda et al. [13] gave constant factor approximation algorithms for several variants of this problem using the local-ratio technique. Partial versions have also been studied for the Facility Location problem and its variants. Charikar et al. [14] explored the outlier or robust version of the uncapacitated facility location problem and k-center problem where only a fraction of the clients need to be serviced. For both the problems they gave constant approximation algorithms. Apart from these there are several other partial covering problems: e.g. k-median with outliers [15], k-MST problem [16] and k-multicut problem [17]. However, these approaches do not seem to work for the Partition-VC problem.

The set of partitions of the edge set in the Partition-VC problem is a special case of matroids. There has been significant work on maximizing a submodular function under matroid constraints [18,19], but none of these results apply to the Partition-VC problem.

The rest of the paper is organized as follows. We present the hardness of the Partition-VC problem in Section 2. Our rounding algorithm and its analysis is presented in Section 3. In Section 4 we present the primal-dual based analysis. Finally concluding remarks are made in Section 5.

2. Hardness of the Partition-VC problem

In this section, we prove that it is NP-hard to get better than $O(\log r)$-approximation for the Partition-VC problem. We give a reduction from the Set Cover problem. Recall that an instance $\mathcal{I}$ of the set cover problem consists of a set X containing r elements, and a set $\mathcal{S}$ of subsets $S_1, \ldots, S_m$ of X. The goal is to find minimum number of sets in $\mathcal{S}$ such that their union is X. This problem is known to be NP-hard, and in fact, it is known that unless P = NP, any polynomial time algorithm for the Set Cover problem must have approximation ratio of $\Omega(\log r)$ [20].

We now describe the reduction in detail. Let $\mathcal{I}$ be an instance of the Set Cover problem as described above. We now construct an instance $\mathcal{I}'$ of the Partition-VC problem. The graph G' in $\mathcal{I}'$ is a bipartite graph. The vertices on left side, V'_L are defined as follows: for each set $S_i \in \mathcal{S}$, we add a vertex s'_i to V'_L. All these vertices have unit cost. The vertices on the right side, V'_R are as follows: for every element $u \in X$, we have a corresponding vertex $u' \in V'_R$. Each of these vertices has infinite cost. Now, we define the set of edges E' and the partition $P_1, \ldots, P_r$ (note that the number of sets in the partition is same as the size of the set X in $\mathcal{I}$). For a vertex $s'_i \in V'_L$ and $u' \in V'_R$, we have an edge between them in E' iff $u \in S_i$ in the instance $\mathcal{I}$. We partition the set E' as follows: for every $u' \in V'_R$, define $P_{u'}$ as the set of edges incident to u'. The partition of E' is $\{P_{u'}\colon\ u' \in V'_R\}$. Further, the quantities $k_{u'}$, which tell how many of the edges in the set $P_{u'}$ need to be covered, are 1. This completes the description of the instance $\mathcal{I}'$. The following lemma is now easy to see.