A Primal-Dual 3-Approximation Algorithm for Hitting 4-Vertex Paths

Eglantine Camby* Jean Cardinal† Mathieu Chapelle† Samuel Fiorini*
Gwenaël Joret‡

Abstract

We consider the problem of removing a minimum number of vertices of a given graph G so that the resulting graph does not contain the path P_k on k vertices as a subgraph. (Thus for $k = 2$ this is the vertex cover problem.) While for $k \in \{2, 3\}$ the problem admits a 2-approximation algorithm, nothing better than a trivial k-approximation is known for $k \geqslant 4$. Our main contribution is a 3-approximation algorithm in the case $k = 4$, that is, for hitting P_4's. The algorithm is inspired by the elegant primal-dual 2-approximation algorithm of Chudak, Goemans, Hochbaum, and Williamson (Operations Research Letters, 1998) for the feedback vertex set problem.

1 Introduction

We consider graphs that are finite, simple, and undirected. Let $\mathscr{F}$ be a family of graphs. In the *$\mathscr{F}$-hitting set problem* we are given a graph G and an integer r in input, and are asked whether at most r vertices can be removed from G such that the resulting graph does not contain a graph from $\mathscr{F}$ as a (nonnecessarily induced) subgraph. This contains as special cases the vertex cover and the feedback vertex set problems, which are obtained by taking $\mathscr{F} = \{P_2\}$ and $\mathscr{F} = \{C_k : k \geqslant 3\}$, respectively, where P_k (reps. C_k) denotes the path (reps. cycle) on k vertices.

We consider the $\mathscr{F}$-hitting set problem in the case where $\mathscr{F}$ consists of the path P_k on $k \geqslant 2$ vertices. As is well known, the vertex cover problem admits a 2-approximation algorithm, and it is widely believed that no better approximation ratio can be achieved unless P = NP (something which is known to hold assuming the Unique Games Conjecture [5]). A straightforward reduction shows that any α-approximation algorithm for the P_k-hitting set problem can be turned into an α-approximation algorithm for the vertex cover problem. Hence, obtaining such an algorithm with $\alpha < 2$ for some $k \geqslant 2$ is unlikely.

For $k = 3$ at least, a 2-approximation can be achieved in polynomial time: As observed in [7], the primal-dual 2-approximation algorithm of Chudak, Goemans, Hochbaum and Williamson [3] for the feedback vertex set problem readily gives a 2-approximation for this problem when adapted in the obvious manner. At a high level point of view, this is essentially due to the fact that graphs not containing P_3 as a subgraph are very restricted forests (every component being isomorphic to K_1 or K_2). However, for $k \geqslant 4$, graphs containing no P_k as a subgraph may have cycles, and the algorithm of Chudak *et al.* no longer yields an approximation algorithm for the problem of hitting P_k's.

While the P_k-hitting set problem has been studied previously [1, 2, 4, 6, 7], nothing better than a k-approximation is known for the problem when $k \geqslant 4$. Note that a k-approximation can trivially be obtained by taking all vertices in an inclusion-wise maximal packing of vertex-disjoint subgraphs each isomorphic to P_k. This motivates the following question, which forms the basis of our investigations: Is there an $\varepsilon > 0$ such that the P_k-hitting set problem can be approximated to within a factor of $(1 - \varepsilon)k$ in polynomial time for every fixed $k \geqslant 3$? In this abstract we report on a first result in that direction,

[1]Université Libre de Bruxelles, Département de Mathématique, CP 216, Boulevard du Triomphe, 1050 Brussels, Belgium, `{ecamby,sfiorini}@ulb.ac.be`
[2]Université Libre de Bruxelles, Département d'Informatique, CP 212, Boulevard du Triomphe, 1050 Brussels, Belgium, `{jcardin,mathieu.chapelle}@ulb.ac.be`
[3]The University of Melbourne, Department of Mathematics and Statistics, 240 Grattan street, Parkville, Melbourne, VIC 3010, Australia, `gwenael.joret@unimelb.edu.au`

namely that there exists a 3-approximation algorithm in the case $k = 4$. Our algorithm follows a primal-dual approach and is very much inspired by the 2-approximation of Chudak *et al.* for the feedback vertex set problem mentioned above.

2 An Approximation Algorithm

In this section we describe the approximation algorithm, Algorithm 1. It can be shown that it approximates the P_4-hitting set problem to within a factor of 3:

Theorem 1. *Algorithm 1 is a 3-approximation algorithm for the P_4-hitting set problem.*

Algorithm 1 Primal-dual 3-approximation algorithm

Require: G a graph
Ensure: X a P_4-hitting set of G
  $X \leftarrow \emptyset$, $G' \leftarrow G$, $\ell \leftarrow 0$
  $y_S \leftarrow 0$ for every subset $S \subseteq V(G)$
  $z_P \leftarrow 0$ for every subgraph P of G isomorphic to P_4
  while G' contains a subgraph isomorphic to P_4 **do**
    $\ell \leftarrow \ell + 1$
    $V' \leftarrow V(G')$
    if G' contains a *good* P_4, say P, **then**
      Increase z_P until $\exists v_\ell \in V' : \sum_{S:v_\ell \in S}(d_S(v_\ell) - 1)y_S + \sum_{P:v_\ell \in V(P)} z_P = 1$
    else
      Increase $y_{V'}$ until $\exists v_\ell \in V' : \sum_{S:v_\ell \in S}(d_S(v_\ell) - 1)y_S + \sum_{P:v_\ell \in V(P)} z_P = 1$
    end if
    $X \leftarrow X \cup \{v_\ell\}$
    $G' \leftarrow G' - v_\ell$
    Remove from G' every component isomorphic to a triangle or a star
  end while
  for $j \leftarrow \ell$ downto 1 **do**
    if $X - \{v_j\}$ is a P_4-hitting set **then**
      $X \leftarrow X - \{v_j\}$
    end if
  end for
  return X

In Algorithm 1, a path P isomorphic to P_4 in a graph H is said to be a *good* P_4 *in* H if for some vertex $v \in V(P)$, every subgraph of H isomorphic to P_4 that includes v also includes another vertex from P. Note that this can be tested polynomial time. Observe also that every inclusion-wise minimal P_4-hitting set of H avoids the vertex v, and thus contains at most 3 vertices from P.

The variables y_S ($S \subseteq V(G)$) and z_P (P subgraph of G isomorphic to P_4) altogether form a solution to the dual of the LP relaxation of the following IP formulation of the problem:

$$\begin{array}{lll}
\textbf{MIN} & \displaystyle\sum_{v \in V(G)} x_v & \\
\text{s.t.} & \displaystyle\sum_{v \in V(P)} x_v \geqslant 1 & \forall P \in \mathcal{P}_4(G); \\
 & \displaystyle\sum_{v \in S}(d_S(v) - 1)x_v \geqslant |E(G[S])| - |S| & \forall S \subseteq V(G); \\
 & x_v \in \{0,1\} & \forall v \in V(G).
\end{array}$$

Above, $\mathcal{P}_4(G)$ denotes the set of subgraphs $P \subseteq G$ that are isomorphic to P_4, and $d_S(v)$ denotes the degree of vertex v in $G[S]$. The first set of inequalities simply encode the fact that each P_4 must be hit at least once. The second set of inequalities are 'sparsity inequalities' similar to those used in [3] for the feedback vertex set problem. To see that they are valid, first consider the case $S = V(G)$: If $X \subseteq V(G)$ is a P_4-hitting set of G then letting $V = V(G)$ we have

$$\begin{aligned}
\sum_{v \in X} (d_V(v) - 1) &= \sum_{v \in X} d_V(v) - |X| \\
&= 2|E(G[X])| + |E(X, V - X)| - |X| \\
&= |E(G)| + |E(G[X])| - |E(G[V - X])| - |X| \\
&= |E(G)| - |V| + |E(G[X])| - |E(G[V - X])| + |V - X| \\
&\geqslant |E(G)| - |V| - |E(G[V - X])| + |V - X| \\
&\geqslant |E(G)| - |V|.
\end{aligned}$$

The last inequality is the key inequality here. It follows from the fact that $|E(H)| \leqslant |V(H)|$ for every graph H not containing a P_4 subgraph, as is easily checked.

Finally, the validity in the case of an arbitrary subset S of $V(G)$ follows from the observation that if X is a P_4-hitting set of G then $X \cap S$ a P_4-hitting set of $G[S]$.

In the algorithm we make sure that the dual solution remains feasible at every step, and use it to choose which vertex to include in the hitting set at each step. The crucial inequality to satisfy for the dual solution is

$$\sum_{S : v \in S} (d_S(v) - 1) y_S + \sum_{P : v \in V(P)} z_P \leqslant 1$$

for every $v \in V(G)$.

At the end of the while-loop we remove from G' the components that do not contain any copy of P_4. These are the components that are isomorphic to a star or a triangle. (Note that K_1 is considered to be a star.)

The last step of the algorithm is a classical 'reverse-delete' step and ensures a certain form of minimality for the P_4-hitting set X output by the algorithm, which is needed in the analysis.

We note that, as is usual with primal-dual algorithms, our 3-approximation algorithm also works for the vertex-weighted version of the problem.

References

[1] Boštjan Brešar, Marko Jakovac, Ján Katrenič, Gabriel Semanišin, and Andrej Taranenko, *On the vertex k-path cover*, Discrete Appl. Math. **161** (2013), no. 13-14, 1943–1949. MR 3056999

[2] Boštjan Brešar, František Kardoš, Ján Katrenič, and Gabriel Semanišin, *Minimum k-path vertex cover*, Discrete Applied Mathematics **159** (2011), no. 12, 1189–1195.

[3] Fabián A Chudak, Michel X Goemans, Dorit S Hochbaum, and David P Williamson, *A primal–dual interpretation of two 2-approximation algorithms for the feedback vertex set problem in undirected graphs*, Operations Research Letters **22** (1998), no. 4, 111–118.

[4] František Kardoš, Ján Katrenič, and Ingo Schiermeyer, *On computing the minimum 3-path vertex cover and dissociation number of graphs*, Theoret. Comput. Sci. **412** (2011), no. 50, 7009–7017. MR 2894368 (2012h:05327)

[5] Subhash Khot and Oded Regev, *Vertex cover might be hard to approximate to within* $2-\epsilon$, J. Comput. Syst. Sci. **74** (2008), no. 3, 335–349.

[6] Jianhua Tu and Wenli Zhou, *A factor 2 approximation algorithm for the vertex cover P_3 problem*, Inform. Process. Lett. **111** (2011), no. 14, 683–686. MR 2840533 (2012g:68293)

[7] ———, *A primal–dual approximation algorithm for the vertex cover P_3 problem*, Theoretical Computer Science **412** (2011), no. 50, 7044–7048.