



**HAL**  
open science

## Minimum k-path vertex cover

Boštjan Brešar, František Kardoš, Ján Katrenič, Gabriel Semanišin

► **To cite this version:**

Boštjan Brešar, František Kardoš, Ján Katrenič, Gabriel Semanišin. Minimum k-path vertex cover. *Discrete Applied Mathematics*, 2011, 159 (12), pp.1189-1195. 10.1016/j.dam.2011.04.008 . inria-00634981

**HAL Id: inria-00634981**

**<https://inria.hal.science/inria-00634981>**

Submitted on 10 Nov 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Minimum $k$ -path vertex cover

Boštjan Brešar<sup>a,b</sup>, František Kardoš<sup>c</sup>, Ján Katrenič<sup>d</sup>, Gabriel Semanišin<sup>d</sup>

<sup>a</sup>*Faculty of Natural Science and Mathematics, University of Maribor, Slovenia*

<sup>b</sup>*Institute of Mathematics, Physics and Mechanics, Ljubljana, Slovenia*

<sup>c</sup>*Institute of Mathematics, P.J. Šafárik University, Košice, Slovakia*

<sup>d</sup>*Institute of Computer Science, P.J. Šafárik University, Košice, Slovakia*

---

## Abstract

A subset  $S$  of vertices of a graph  $G$  is called a  $k$ -path vertex cover if every path of order  $k$  in  $G$  contains at least one vertex from  $S$ . Denote by  $\psi_k(G)$  the minimum cardinality of a  $k$ -path vertex cover in  $G$ . It is shown that the problem of determining  $\psi_k(G)$  is NP-hard for each  $k \geq 2$ , while for trees the problem can be solved in linear time. We investigate upper bounds on the value of  $\psi_k(G)$  and provide several estimations and exact values of  $\psi_k(G)$ . We also prove that  $\psi_3(G) \leq (2n + m)/6$ , for every graph  $G$  with  $n$  vertices and  $m$  edges.

---

## 1. Introduction and motivation

In this paper we consider finite graphs without loops and multiple edges and use standard graph theory notations (see e.g. [30]). In particular, by the *order of a path*  $P$  we understand the number of vertices on  $P$  while the *length of a path* is the number of edges of  $P$ .

We introduce a new graph invariant that generalizes the intensively studied concept of vertex cover. Our research is motivated by the following problem [25] related to secure communication in wireless sensor networks (shortly WSN). The topology of WSN can be represented by a graph, in which vertices represent sensor devices and edges represent communication channels between pairs of sensor devices. Traditional security techniques cannot be applied directly to WSN, because sensor devices are limited in their

---

*Email addresses:* `bostjan.bresar@uni-mb.si` (Boštjan Brešar),  
`frantisek.kardos@upjs.sk` (František Kardoš), `jan.katrenic@upjs.sk` (Ján  
Katrenič), `gabriel.semanisin@upjs.sk` (Gabriel Semanišin)

computation, energy, communication capabilities. Moreover, they are often deployed in accessible areas, where they can be captured by an attacker. In general, a standard sensor device is not considered as tamper-resistant and it is undesirable to make all devices of a sensor network tamper-proof due to increasing cost. Therefore, the design of WSN security protocols has become a challenge in security research.

We focus on the Canvas scheme [13, 28, 25, 26] which should provide data integrity in a sensor network. The  $k$ -generalized Canvas scheme [25] guarantees data integrity under the assumption that at least one node which is not captured exists on each path of the length  $k - 1$  in the communication graph. The scheme combines the properties of cryptographic primitives and the network topology. The model distinguishes between two kinds of sensor devices – *protected* and *unprotected*. The attacker is unable to copy secrets from a protected device. This property can be realized by making the protected device tamper-resistant or placing the protected device on a safe location, where capture is problematic. On the other hand, an unprotected device can be captured by the attacker, who can also copy secrets from the device and gain control over it. During the deployment and initialization of a sensor network, it should be ensured, that at least one protected node exists on each path of the length  $k - 1$  in the communication graph [25]. The problem to minimize the cost of the network by minimizing the number of protected vertices is formulated in [25].

Formally, let  $G$  be a graph and let  $k$  be a positive integer. A subset of vertices  $S \subseteq V(G)$  is called a  $k$ -path vertex cover if every path of order  $k$  in  $G$  contains at least one vertex from  $S$ . We denote by  $\psi_k(G)$  the minimum cardinality of a  $k$ -path vertex cover in  $G$ .

Clearly, 2-path vertex cover corresponds to the well know vertex cover (a subset of vertices such that each edge of the graph is incident to at least one vertex of the set). Therefore  $\psi_2(G)$  is equal to the size of the minimum vertex cover of a graph  $G$ . Moreover, the value of  $\psi_3(G)$  corresponds to the so-called dissociation number of a graph [5, 14, 32] defined as follows. A subset of vertices in a graph  $G$  is called a *dissociation set* if it induces a subgraph with maximum degree 1. The number of vertices in a maximum cardinality dissociation set in  $G$  is called the *dissociation number* of  $G$  and is denoted by  $diss(G)$ . Clearly,  $\psi_3(G) = |V(G)| - diss(G)$ .

A related colouring is the  $k$ -path chromatic number, which is the minimum number of colours that are necessary for colouring the vertices of  $G$  in such a way that each colour class forms a  $P_k$ -free set [1, 12, 20, 19, 24], and it is

easy to see that  $\psi_k(G) \leq \frac{\chi_k(G)-1}{\chi_k(G)}|V(G)|$ .

Since the minimum vertex cover problem is NP-hard [21], it is not surprising that so is the problem of determining  $\psi_k$  for each  $k \geq 3$ . We provide details in Section 2 – actually, we reduce the minimum vertex cover problem to the minimum  $k$ -path vertex cover problem.

However, since the question whether there is a  $k$ -path vertex cover of size at most  $t$  can be expressed in the monadic second order logic, by famous Courcelle’s theorem [8], the minimum  $k$ -path vertex cover problem can be solved in linear time on graphs with bounded treewidth, e.g. trees, series-parallel graphs, outerplanar graphs, etc. In Section 3 we determine the exact value of  $\psi_k$  for trees and present a linear time algorithm which returns an optimal solution for trees. Then Section 4 is devoted to outerplanar graphs. We present a tight upper bound for  $\psi_3(G)$ .

In Section 5 we provide several estimations for the size of minimum  $k$ -path vertex cover on degree of its vertices. Finally, we prove that  $\psi_3(G) \leq (2n + m)/6$ , for every graph  $G$  with  $n$  vertices and  $m$  edges.

## 2. NP-completeness

In this section we prove that the problem of determining  $\psi_k$  is NP-complete for every  $k \geq 2$ . For  $k = 2$  this equals to the vertex cover problem (shortly VCP) which is known to be NP-complete [21]. For an arbitrary fixed integer  $k \geq 2$  the decision version of the  $k$ -path vertex cover problem can be stated as follows:

**$k$ -Path Vertex Cover Problem ( $k$ -PVCP)**

INSTANCE: Graph  $G$  and a positive integer  $t$ .

QUESTION: Is there a  $k$ -path vertex cover  $S$  for  $G$  of size at most  $t$ ?

**Theorem 1.** *For any fixed integer  $k \geq 2$  the  $k$ -Path Vertex Cover Problem is NP-complete.*

*Proof.* First, it is not hard to see that  $k$ -PVCP is in NP. Note that there are at most  $O(n^k)$  paths of length  $k - 1$  in  $G$ . Thus for a given set  $S$  we can decide in polynomial time whether all paths of length  $k - 1$  are secured by the vertices from  $S$ .

Note that 2-PVCP coincides with VCP. For  $k > 2$  we will reduce VCP to  $k$ -PVCP. Let  $G$  be an arbitrary graph. Let  $G'$  be a graph obtained from

$G$  by joining a path with  $\lfloor \frac{k-1}{2} \rfloor$  new vertices to every vertex of  $G$  (i.e. for  $k = 3$  or  $k = 4$  we attach a leaf to each vertex; for  $k = 5$  or  $k = 6$  a path on two vertices is attached to each vertex of  $G$ ; etc.). We call the vertices of  $G$  *original vertices*, and others are called *new*. In order to finish the proof it suffices to show that the size of a minimum vertex cover in  $G$  is equal to  $\psi_k(G')$ .

Suppose that  $G'$  has a  $k$ -path vertex cover  $S'$ . Suppose that  $S'$  contains also a new vertex  $x$  (i.e. it lies on an attached path). Let  $x'$  be an original vertex that has the smallest distance from  $x$ . Note that  $x$  secures only the paths that go through the vertex  $x'$ , since the length of the attached path is less than  $k$ . Hence, if  $x$  is replaced in  $S'$  by  $x'$ , all paths remain secured. By consecutive application of this procedure for each new vertex  $x$  we obtain a set  $S$  comprising only of the original vertices. Clearly,  $|S| \leq |S'|$ , and by the procedure described above,  $S$  is also a  $k$ -path vertex cover. We claim that  $S \subseteq V(G)$  forms a vertex cover in  $G$ .

Suppose to the contrary that there is an edge  $xy$  in  $G$  whose endvertices  $x$  and  $y$  are not in  $S$ . Consider the path  $P$  in  $G'$ , composed by the path, attached to  $x$ , vertices  $x$  and  $y$ , and the path attached to  $y$ . Then clearly  $P$  contains no vertex from  $S$ , and it has  $2\lfloor \frac{k-1}{2} \rfloor + 2 \geq k$  vertices, which is a contradiction. Thus  $S$  is a vertex cover in  $G$  of size at most  $|S'|$ .

Conversely, by a similar reasoning we can show that a vertex cover of  $G$  yields a  $k$ -path vertex cover set in  $G'$  of the same size.  $\square$

The reduction described in the proof of Theorem 1 also shows that for any approximation rate  $r$  one can transform a polynomial time  $r$ -approximation for the minimum  $k$ -path vertex cover problem to a polynomial time  $r$ -approximation algorithm for the minimum vertex cover. Using a result of [10], we immediately infer that it is NP-hard to approximate  $k$ -PVCP, for  $k > 2$ , within a factor of 1.3606, unless  $P = NP$ .

A well-known 2-approximation algorithm which repeatedly puts vertices of an edge into the constructed vertex cover and removes them from the graph, was discovered independently by F. Gavril and M. Yannakakis [7]. A polynomial time approximation with a better constant factor is not known for the vertex cover problem. Similarly, for the minimum  $k$ -path vertex cover problem we can construct an approximation algorithm in the following way: Find a path on  $k$  vertices, put its  $k$  vertices into the constructed set and remove them from the graph. Since at least one of these  $k$  vertices belongs to an optimal solution, this provides a  $k$ -approximation algorithm

for the minimum  $k$ -path vertex cover with a polynomial running time while  $k$  is considered as a fixed constant. Note that the currently best known randomized algorithm for finding a path on  $k$  vertices runs in time  $O(2^k n^{O(1)})$ , where  $n$  denotes the order of a graph (see [31]).

### 3. Path vertex cover for trees

We begin our investigation with the class of trees, that present an important underlying communication topology for WSN. Courcelle's theorem [8] guarantees the existence of a linear time algorithm, basically using dynamic programming. In this section, we describe such an algorithm in detail, and we use it then to derive a sharp upper bound  $\psi_k(T) \leq |V(T)|/k$  for an arbitrary tree  $T$ .

In order to simplify our consideration, consider that the input tree is rooted at a vertex  $u$ . By *properly rooted subtree* we denote a subtree  $T_v$ , induced by a vertex  $v$  and its descendants (with respect to  $u$  as the root) and satisfies the following properties:

1.  $T_v$  contains a path on  $k$  vertices;
2.  $T_v \setminus v$  does not contain a path on  $k$  vertices.

The algorithm `PVCPTree` systematically searches for a properly rooted tree  $T_v$ , puts  $v$  into a solution and removes  $T_v$  from the input tree  $T$ .

---

**Function** `PVCPTree`( $T, k$ )

---

**Input:** A tree  $T$  on  $n$  vertices and a positive integer  $k$ ;  
**Output:** A  $k$ -path vertex cover  $S$  of  $T$ ;  
 Form an arbitrary vertex  $u \in T$ , make  $T$  rooted in  $u$ ;  
 $S := \emptyset$ ;  
**while**  $T$  contains a properly rooted subtree  $T_v$  **do**  
      $S := S \cup \{v\}$ ;  
      $T := T \setminus T_v$ ;  
**return**  $S$ ;

---

**Theorem 2.** *Let  $T$  be a tree and  $k$  be a positive integer. The algorithm `PVCPTree`( $T, k$ ) returns an optimal  $k$ -path vertex cover of  $T$  of size at most  $\frac{|V(T)|}{k}$ . Therefore,  $\psi_k(T) \leq \frac{|V(T)|}{k}$ .*

*Proof.* First, we shall argue that `PVCPTree` returns an optimal solution. We prove this by induction on the number of vertices in the tree  $T$ . If  $T$  does

not contain any path on  $k$  vertices, the empty set is the optimal solution. Suppose  $T$  contains a path on  $k$  vertices. Let  $T_v$  be a properly rooted subtree of  $T$ . Since any  $k$ -path cover of  $T$  contains a vertex of  $T_v$  it follows that  $\psi_k(T) = \psi_k(T \setminus T_v) + 1$  and hence the result follows by induction.

To prove that the returning set  $S$  contains at most  $\frac{|V(T)|}{k}$  vertices, we argue that each loop of the algorithm inserts one vertex into  $S$  and removes from  $T$  one properly rooted subtree having at least  $k$  vertices.  $\square$

Concerning the time complexity of the algorithm, it is straightforward to implement the algorithm `PVCPTree` such that the returning set  $S$  is computed in linear time.

#### 4. Outerplanar graphs

In this section we study the 3-path vertex cover of outerplanar graphs.

**Theorem 3.** *Let  $G$  be an outerplanar graph of order  $n$ . Then  $\psi_3(G) \leq \frac{n}{2}$ .*

*Proof.* We prove the statement by induction on the number of vertices of  $G$ .

Let  $H$  be a maximal outerplanar graph such that  $G$  is its subgraph satisfying  $V(H) = V(G)$ . (Recall that  $H$  is *maximal outerplanar* if  $H$  is outerplanar, but adding any edge destroys that property.) It is easy to see that  $H$  is 2-connected and all its inner faces are triangles. Observe that every  $k$ -path vertex cover in  $H$  is a  $k$ -path vertex cover in  $G$ , since  $G$  is a subgraph of  $H$ . Therefore, it suffices to find a 3-path vertex cover in  $H$  of size at most  $\frac{n}{2}$ .

Obviously, if  $H$  consists of a single triangle, then  $\psi_3(H) = 1 \leq \frac{3}{2} = \frac{n}{2}$ .

Assume  $H$  has at least four vertices. Since  $H$  is a 2-connected outerplanar graph, the closed trail bounding the outer face contains each vertex of  $H$  precisely once, hence  $H$  is Hamiltonian. Let  $v_1, v_2, \dots, v_n$  be the cyclic ordering of vertices of  $H$  along the Hamiltonian cycle. Colour a vertex  $v_i$  white, if the degree of  $v_i$  in  $H$  is 2, otherwise colour it black. Since all the inner faces of  $H$  are triangles, there are no two consecutive white vertices, unless  $H$  consists of a single triangle, which is excluded. Hence, the white vertices induce an independent set. The edge  $v_i v_{i+1}$  is called *good*, if it has a white endvertex, otherwise it is *bad*. If all the edges incident with the outer faces are good, it implies that  $n$  is even, and half of the vertices are white. Then the set of all black vertices is a 3-path vertex cover of size  $\frac{n}{2}$ , since the white vertices form an independent set.

Assume there is a bad edge, i.e. there is at least one pair of consecutive black vertices, say  $v_i, v_{i+1}$ . We claim that there is an edge  $v_i v_j$  of  $H$  such that  $v_i, v_{i+1}$ , and  $v_j$  form a triangular face adjacent to the outer face, all the three vertices  $v_i, v_{i+1}$ , and  $v_j$  are black, and all the edges  $v_{i+1}v_{i+2}, v_{i+2}v_{i+3}, \dots, v_{j-1}, v_j$  (indices taken modulo  $n$ ) are good.

For the sake of contradiction, suppose that this is not the case. Let  $e = v_i v_{i+1}$  be a bad edge. There is a unique triangular face of  $H$  incident both with  $v_i$  and  $v_{i+1}$ ; let  $v_j$  be the third vertex incident with the face. It is easy to see that  $v_j$  must be black. The vertices  $v_i, v_{i+1}$ , and  $v_j$  cut the Hamiltonian cycle into the edge  $v_i v_{i+1}$  and two paths; let  $\sigma(e)$  be the smaller of their lengths.

Let  $e_0 = v_{i_0} v_{i_0+1}$  be the bad edge with  $\sigma(e)$  minimal; let  $v_{j_0}$  be the corresponding black vertex such that together with  $v_{i_0}$  and  $v_{i_0+1}$  they form a triangular face of  $H$ . Without loss of generality assume that  $P = v_{i_0+1}, \dots, v_{j_0}$  is the path of length  $\sigma(e)$ . Observe that  $P$  contains at least one white vertex, thus, it contains good edges. If all the edges  $v_{i_0+1}v_{i_0+2}, \dots, v_{j_0-1}v_{j_0}$  are good, we are done. Otherwise, we find a bad edge  $e_1 = v_{i_1} v_{i_1+1}$  of  $P$ . Again, there is a unique triangular face of  $H$  incident both with  $v_{i_1}$  and  $v_{i_1+1}$ ; let  $v_{j_1}$  be the third (black) vertex incident with the face. Since  $\{v_{i_0+1}, v_{j_0}\}$  is a 2-cut of  $H$  separating the vertices  $v_{i_0+2}, \dots, v_{j_0-1}$  from the rest of the graph, we have  $v_{j_1} \in \{v_{i_0+1}, \dots, v_{j_0}\}$ . But then  $\sigma(e_1) < \sigma(e_0)$ , a contradiction with the choice of  $e_0$ .

Therefore, there is an edge  $e = v_i v_j$  of  $H$  such that  $v_i, v_{i+1}$ , and  $v_j$  form a triangular face adjacent to the outer face, all the three vertices  $v_i, v_{i+1}$ , and  $v_j$  are black, and all the edges  $v_{i+1}v_{i+2}, \dots, v_{j-1}v_j$  are good. It means that on the path  $P = v_{i+1}, \dots, v_j$  black and white vertices alternate, thus,  $\sigma(e) = 2s$  for some integer  $s$ .

Let  $W$  be the set of white vertices on  $P$ , let  $B$  be the set of black vertices on  $P$ , including  $v_{i+1}$  and  $v_j$ . Obviously  $|W| = s$  and  $|B| = s+1$ ; vertices from  $W$  induce an independent set in  $H$  and  $v_{i+1}$  is adjacent to precisely one of them. Let  $S'$  be a 3-path vertex cover in the graph  $H' = H \setminus \{v_i, v_{i+1}, \dots, v_j\}$  of size at most  $\frac{|V(G')|}{2} = \frac{n-2s-2}{2}$  given by induction. Then  $S = S' \cup \{v_i\} \cup (B \setminus \{v_{i+1}\})$  is a 3-path vertex cover in  $H$  of size at most  $\frac{n-2s-2}{2} + 1 + s = \frac{n}{2}$ .  $\square$

The bound  $\frac{n}{2}$  in Theorem 3 is the best possible, since it is easy to find outerplanar graphs with  $\psi_3 \geq \frac{n}{2}$ . Consider an arbitrary 2-connected outerplanar graph  $G$ . Let  $v_1, v_2, \dots, v_n$  be the ordering of its vertices along the Hamiltonian cycle (the boundary of the outer face). For each edge  $v_i v_{i+1}$



incident with the outer face ( $v_{n+1} = v_1$ ), add a new vertex  $u_i$  and two new edges  $u_i v_i$  and  $u_i v_{i+1}$ . Let the resulting graph be  $H$ . It is easy to see that  $H$  is a 2-connected outerplanar graph on  $2n$  vertices. We claim that  $\psi_3(H) \geq n$ . Let  $S$  be a 3-path vertex cover in  $H$ . Divide the vertices of  $H$  into  $n$  pairs of the form  $v_i, u_i$ . Observe that if  $u_i \notin S$  and  $v_i \notin S$ , then  $u_{i-1} \in S$  and  $v_{i-1} \in S$  (otherwise there is a path on 3 vertices in  $H$  not covered by  $S$ ). Therefore, the average number of vertices in  $S$  is either at least  $\frac{1}{2}$  (if at least one vertex of a pair is in  $S$ ), or  $\frac{2}{4}$  (if no vertex is in  $S$ , then in the previous pair both are). Altogether,  $|S| \geq n$ .

## 5. Upper bounds on degree of vertices

In this section we provide several upper bounds on path vertex cover based on degrees of a graph. First, recall theorem of Caro [6] and Wei [29], which states

$$\psi_2(G) \leq |V(G)| - \sum_{u \in V(G)} \frac{1}{1 + d(u)}$$

for every graph  $G$ . Since the only graphs for which this is best-possible are the disjoint unions of cliques, additional structural assumptions allow improvements (see e.g. [15, 16, 17, 18, 23]).

The problem of dissociation number of graph was also studied in [14], where Göring et al. proved

$$\psi_3(G) \leq |V(G)| - \sum_{u \in G} \frac{1}{1 + d(u)} - \sum_{uv \in E(G)} \frac{2}{|N(u) \cup N(v)|(|N(u) \cup N(v)| - 1)}.$$

In what follows, we provide the following generalization of Caro-Wei theorem for  $\psi_k(G)$ .

**Theorem 4.** *For every graph  $G$  without isolated vertices:*

$$\psi_k(G) \leq |V(G)| - \frac{k-1}{k} \sum_{u \in V(G)} \frac{2}{1 + d(u)}.$$

*Proof.* Let us arbitrarily order the vertices of  $G$ , and let us start with  $S$  being the empty set. One by one let us add vertex  $v_i$  to  $S$  unless at least two of its neighbours are already there. The probability that  $v_i$  will eventually land in  $S$ , is  $\frac{2}{1+d(v_i)}$ , because it is the probability that in random ordering of

vertices of  $G$ ,  $v_i$  precedes each of its neighbours except for one. From this one can deduce that the expected size of the set  $S$  is at least  $\sum_{v_i \in V(G)} \frac{2}{(1+d(v_i))}$ . Since  $S$  is a 1-degenerated graph,  $S$  is a forest. Using Theorem 2 we get  $\psi_k(S) \leq \frac{1}{k}|V(S)|$ . Finally, to construct a  $k$ -path vertex cover for  $G$  we put into solution  $V(G) \setminus S$  and all the vertices forming the minimum  $k$ -path vertex cover of  $S$ .  $\square$

In the rest of this section, we investigate the upper bounds to the 3-path vertex cover based on maximum and average degree of a graph. First, recall a well-known decomposition theorem of Lovász [22]:

**Theorem 5.** *If  $s$  and  $t$  are non-negative integers, and if  $G$  is a graph with maximum degree at most  $s+t+1$ , then the vertex set of  $G$  can be partitioned into two sets which induce subgraphs of maximum degree at most  $s$  and  $t$ , respectively.*

The following is a straightforward generalization of Theorem 5 for  $k$ -sets (see e.g. [9]).

**Theorem 6.** *For any graph  $G$  of maximum degree  $\Delta$ , the vertex set of  $G$  can be partitioned into  $k$  sets which induce subgraphs of maximum degree at most  $\frac{\Delta}{k}$ . Moreover, this can be computed in running time  $O(\Delta|E(G)|)$ .*

**Corollary 7.** *Let  $G$  be a graph of maximum degree  $\Delta$ . Then*

$$\psi_3(G) \leq \frac{\lceil \frac{\Delta-1}{2} \rceil}{\lceil \frac{\Delta+1}{2} \rceil} |V(G)|.$$

*Moreover, such a solution can be computed in running time  $O(\Delta|E(G)|)$ .*

*Proof.* According to Theorem 6, the vertex set of  $G$  can be partitioned into  $\lceil \frac{\Delta+1}{2} \rceil$  sets which induce subgraphs of maximum degree at most 1. Let  $S$  denote the largest set of such a partition; clearly  $|S| \geq \frac{|V(G)|}{\lceil \frac{\Delta+1}{2} \rceil}$ . To create a 3-path vertex cover set for  $G$ , we protect  $V(G) \setminus S$ .  $\square$

**Corollary 8.** *There is a linear time algorithm which returns a 3-path vertex cover of size at most  $\min(\frac{|V(G)|}{2}, \frac{|E(G)|}{2})$ , for arbitrary graph  $G$  of maximum degree at most 3.*

*Proof.* To construct a 3-path vertex cover of size at most  $|E(G)|/2$  one can repeatedly put into a solution a vertex of degree at least 2. To construct a 3-path vertex cover of size at most  $|V(G)|/2$  for a graph  $G$ ,  $\Delta(G) \leq 3$ , one can use Corollary 7.  $\square$

Next we provide an algorithm which guarantees a tight upper-bound for  $\psi_3$  based on the number of vertices and edges of a graph.

---

**Algorithm 2:** SPARSE<sub>3</sub>( $G$ )

---

**Input:** A graph  $G$  on  $n$  vertices;

**Output:** 3-path vertex cover  $H$ ,  $H \subseteq V(G)$ ;

$H := \emptyset$ ;

**while**  $G$  contains any vertex  $v$  of degree at least 4 **do**

$H := H \cup \{v\}$ ;

Remove from  $G$  the vertex  $v$  and all edges incident with  $v$ ;

$H := H \cup \text{Solve\_SubCubic}(G)$ ;

**return**  $H$ ;

---

Here,  $\text{Solve\_SubCubic}(G)$  denotes the algorithm of Corollary 8.

**Theorem 9.** *Let  $G$  be a graph having  $n$  vertices and  $m$  edges. The algorithm SPARSE<sub>3</sub>( $G$ ) returns a 3-path vertex cover  $H$  of size at most  $\frac{2n+m}{6}$ .*

*Proof.* The algorithm consists of two phases. The first phase repeatedly erases vertices of degree at least 4. After this step, we get a graph  $G$  of maximum degree 3. In the second phase the algorithm uses the method  $\text{Solve\_SubCubic}(G)$ .

Let  $s$  denote the number of vertices (of degree at least 4) protected in the first phase, let  $d$  denote the number of vertices protected in the second phase. From Corollary 8 we get  $s + 2d \leq n$  and  $4s + 2d \leq m$ . Therefore, we get two upper bounds for  $|H|$ .

$$\begin{aligned} |H| &= s + d \leq s + \frac{n-s}{2} = \frac{n}{2} + \frac{s}{2} \\ |H| &= s + d \leq s + \frac{m-4s}{2} = \frac{m}{2} - s \end{aligned}$$

This implies  $3|H| \leq n + \frac{m}{2}$ .  $\square$

**Corollary 10.** *Let  $G$  be a graph on  $n$  vertices and  $m$  edges. Then*

$$\psi_3(G) \leq \frac{2n + m}{6}.$$

**Theorem 11.** *Let  $a, b$  be integers such that  $b \leq a \leq 2b$ . There is a graph on  $n$  vertices and  $m$  edges such that  $\frac{m}{n} = \frac{a}{b}$  and*

$$\psi_3(G) \geq \frac{2n + m}{6}.$$

*Proof.* Let  $x = 3(2b - a)$  and  $y = 2(a - b)$ . We construct the graph  $G$  having two types of components:

- $x$  components  $C_4$  (a cycle on 4 vertices) with 4 edges,  $\psi_3(C_4) = 2$ .
- $y$  components  $H_6$  ( $K_6$  with a perfect matching removed) with 12 edges,  $\psi_3(H_6) = 4$ .

It is easy to see that  $n = 4x + 6y$  and  $m = 4x + 12y$ . Then

$$\frac{2n + m}{6} = \frac{12x + 24y}{6} = 2x + 4y = \psi_3(G).$$

Moreover,

$$\frac{m}{n} = \frac{2x + 6y}{2x + 3y} = \frac{6(2b - a) + 12(a - b)}{6(2b - a) + 6(a - b)} = \frac{6a}{6b} = \frac{a}{b}.$$

□

## 6. Concluding remarks

The  $k$ -path vertex cover of a graph  $G$  can be easily translated to the transversal of the hypergraph with  $V(G)$  as its vertex set whose edges are all the paths of order  $k$  in  $G$ . Since the corresponding hypergraph is  $k$ -uniform, one can infer from the general upper bound of Alon [3] that  $\psi_k(G) \leq |V(G)|p(G) \ln(k)/k$ , where  $p(G)$  is the number of paths of order  $k$  in  $G$ . More specifically, for the case of 3-uniform hypergraphs, the bound of Thomassé and Yeo [27] implies that  $\psi_3(G) \leq (|V(G)| + p(G))/4$ . As expected, the latter general result is only in some special cases better as our bound from Theorem 9, in particular such case is when  $G$  is  $K_3$ -free and  $\sum_{v \in V(G)} \frac{d(v)(d(v)-1)}{2} - 2m < \frac{n}{3}$ .

Yet another approach to the  $k$ -path vertex covering problem is via the concept of *decycling number*  $\nabla(G)$  of a graph  $G$ ; it is defined as the size of a minimum set  $S \subseteq V(G)$  such that  $G - S$  is a forest (such a set  $S$  is

also called a *feedback vertex set* - we refer the reader to [11] for a survey of this parameter). By removing the corresponding subset  $S$  with  $\nabla = \nabla(G)$  vertices from a graph  $G$ , we are left with a forest, and by using the algorithm `PVCPTree` on  $G - S$  we infer by Theorem 2 that  $\psi_k(G) \leq \nabla + \frac{n-\nabla}{k} = \frac{n+(k-1)\nabla}{k}$ . If  $G$  is outerplanar, the bound from [4],  $\nabla(G) \leq \lceil \frac{n}{3} \rceil$  implies roughly  $\psi_3(G) \leq \frac{5}{9}n$  which is worse than the bound  $\psi_3(G) \leq \frac{n}{2}$ , obtained in Theorem 4. Albertson and Berman [2] conjectured that every planar graph has an induced acyclic subgraph with at least half the vertices and this problem is still open. An interesting question remains to prove the bound  $\psi_3(G) \leq \frac{2n}{3}$  for planar graphs, which is a weaker form of the previous conjecture.

## 7. Acknowledgements

The authors would like to thank Yair Caro and two anonymous referees for several ideas and approaches that are briefly mentioned in the last section.

The first author was supported in part by the Ministry of Science and Technology of Slovenia under the grants J1-2043 and P1-0297. The research of the second, third and fourth authors was also supported by the Slovak Research and Development Agency under contracts APVV-0023-10, APVV-0035-10, APVV SK-SI-0014-10 and by the Agency of the Slovak Ministry of Education for the Structural Funds of the EU, under project ITMS:26220120007.

## References

- [1] J. Akiyama, H. Era, S. V. Gervacio, M. Watanabe, Path chromatic numbers of graphs. *J. Graph Theory* 13 (1989) 571–573.
- [2] M. O. Albertson, D. M. Berman, The acyclic chromatic number. *Proceedings of the Seventh Southeastern Conference on Combinatorics, Graph Theory, and Computing, Congr. Numer. 17, Utilitas Math., (1976) 51–69.*
- [3] N. Alon, Transversal numbers of uniform hypergraphs, *Graphs Combin.* 6 (1990) 1–4.
- [4] S. Bau, L. W. Beineke and R. C. Vandell, Decycling snakes, *Congr. Numer.* 134 (1998) 79–87.

- [5] R. Boliac, K. Cameron, V.V. Lozin, On computing the dissociation number and the induced matching number of bipartite graphs, *Ars Comb.* 72 (2004) 241–253.
- [6] Y. Caro, New Results on the Independence Number, Technical Report, Tel-Aviv University, 1979.
- [7] T. Cormen, C.E. Leiserson, R.L. Rivest: Introduction to algorithms. The MIT Press 1990.
- [8] B. Courcelle, The monadic second-order logic of graphs. I. recognizable sets of finite graphs, *Information and Computation* 85 (1990) 12–75.
- [9] L.J. Cowen, C.E. Jesurum, Coloring with defects, DIMACS Technical Report (1995) 95–25.
- [10] I. Dinur, S. Safra, On the hardness of approximating minimum vertex cover, *Annals of Mathematics* 162 (2005) 439–485.
- [11] P. Festa, P. M. Pardalos, M. G. C. Resende, Feedback set problems, AT&T Labs. Research Technical Report TR 99.2.2 (1999), *Encyclopedia of Optimization*, Kluwer Academic Publisher, Dordrecht, Netherlands, (2000).
- [12] M. Frick, F. Bullock, *Detour chromatic number*, *Discuss. Math. Graph Theory* 21 (2001) 283–291.
- [13] D. Gollmann: Protocol analysis for concrete environments. EUROCAST 2005. LNCS, vol. 3643, pp. 365–372. Springer, Heidelberg (2005).
- [14] F. Göring, J. Harant, D. Rautenbach and I. Schiermeyer, On F-independence in graphs, *Discussiones Mathematicae Graph Theory* 29 2 (2009) 377-383.
- [15] J. Harant, M.A. Henning, D. Rautenbach, and I. Schiermeyer, Independence Number in Graphs of Maximum Degree Three, *Discrete Math.* 308 (2008) 5829–5833.
- [16] J. Harant and I. Schiermeyer, On the independence number of a graph in terms of order and size, *Discrete Math.* 232 (2001) 131-138.

- [17] C.C. Heckman, On the Tightness of the  $5/14$  Independence Ratio, *Discrete Math.* 308 (2008) 3169–3179.
- [18] C.C. Heckman and R. Thomas, A New Proof of the Independence Ratio of Triangle-Free Cubic Graphs, *Discrete Math.* 233 (2001) 233–237.
- [19] Y. Jinjiang, Computational complexity of  $(2,2)$  path chromatic number problem. *Appl. Math. J. Chinese Univ. Ser. B.* 10B (1995) 83–88.
- [20] G. Johns, F. Saba, On the path-chromatic number of a graph. *Ann. New York Acad. Sci.* 576 (1989) 275–280.
- [21] R.M. Karp, Reducibility among combinatorial problems, *Complexity of computer computations*, R.E. Miller and J.W. Thatcher (Editors), Plenum Press, New York 1972, 85–104.
- [22] L. Lovász. *On decompositions of graphs*, *Studia Sci. Math Hungar.* 1 (1966) 237–238.
- [23] C. Löwenstein, A. S. Pedersen, D. Rautenbach, F. Regen, Independence, odd girth, and average degree. *Journal of Graph Theory*, n/a. doi: 10.1002/jgt.20518.
- [24] C. M. Mynhardt, I. Broere, Generalized colorings of graphs, *Graph theory with applications to algorithms and computer science*, John Wiley & Sons, Inc., New York, NY, 1985, 583–594.
- [25] M. Novotný, *Design and Analysis of a Generalized Canvas Protocol*. Proceedings of WISTP 2010, LNCS 6033 (2010), Springer-Verlag, 106–121.
- [26] M. Novotný. Formal analysis of security protocols for wireless sensor networks, *Tatra Mt. Math. Publ.* 47 (2010), 81–97.
- [27] S. Thomassé, A. Yeo, Total domination of graphs and small transversals of hypergraphs, *Combinatorica* 27 (2007) 473–487.
- [28] H. Vogt: Integrity preservation for communication in sensor networks. Technical Report 434, Institute for Pervasive Computing. ETH Zürich (2004)

- [29] V.K. Wei, A Lower Bound on the Stability Number of a Simple Graph, Technical memorandum, TM 81 - 11217 - 9, Bell laboratories, 1981.
- [30] D.B. West, Introduction to graph theory, Prentice Hall 1996.
- [31] R. Williams, Finding paths of length  $k$  in  $O^*(2^k)$  time, Inf. Process. Lett. 109 (2009) 315–318.
- [32] M. Yannakakis, Node-deletion problems on bipartite graphs, SIAM J. Computing 10 (1981), 310–327.