On the Path Sequence of a Graph

SŁAWOMIR BAKALARSKI, JAKUB ZYGADŁO
Institute of Computer Science and Computational Mathematics
Faculty of Mathematics and Computer Science, Jagiellonian University
ul. Łojasiewicza 6, 30-348 Kraków, Poland
e-mail: *Slawomir.Bakalarski@uj.edu.pl, Jakub.Zygadlo@ii.uj.edu.pl*

Abstract. A subset S of vertices of a graph $G = (V, E)$ is called a k-path vertex cover if every path on k vertices 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 and form a sequence $\psi(G) = (\psi_1(G), \psi_2(G), \ldots, \psi_{|V|}(G))$, called the path sequence of G. In this paper we prove necessary and sufficient conditions for two integers to appear on fixed positions in $\psi(G)$. A complete list of all possible path sequences (with multiplicities) for small connected graphs is also given.

Keywords: k-path vertex cover, path sequence, list for small graphs.

1. Introduction

Let G be a graph and let k be a positive integer. Following [3], define a *k-path vertex cover* (*k*-PVC for short) of G as a subset S of vertices of G such that every path on k vertices in G has at least one vertex in common with S. A k-PVC is called *minimum* if it has minimum cardinality among all k-path vertex covers of G. This minimum cardinality is denoted by $\psi_k(G)$ and called a *k-path number* of G. Our main subject is to investigate the properties of *path sequences*, defined as follows:

Definition 1 *Let $G = (V, E)$ be a graph on n vertices. The sequence of all path numbers, namely*

$$\psi(G) = (\psi_1(G), \psi_2(G), \ldots, \psi_n(G))$$

will be called a path sequence of G.

The problem of determining path numbers appears to have its motivation in the Canvas protocol for wireless sensor networks [15]. Other applications emerge in the context of security of computer networks [1] and processing of large graphs representing road networks (eg. road planning), see [9]. In all these areas in addition to determining k-PVCs and $\psi_k(G)$ for a fixed k, it is also natural to ask another question: given a number t of vertices (representing tamper-proof sensors, firewalls in a network, POIs on a map, etc.) find path lengths for which t is sufficient in terms of system security, map accuracy, etc. In other words, find the largest number k such that $\psi_k(G) > t$. One can also consider 'upgrades' of the system in question, for example: if we want to decrease k (i.e. increase system security, increase map accuracy, etc.), how much needs t be increased? These questions are clearly related to a path sequence and its subsequences.

Path numbers generalize some well-known problems from graph theory, for example the cardinality of a minimum vertex cover of a graph G equals $\psi_2(G)$, dissociation number of G is equal to $|V| - \psi_3(G)$ (see [3], [12]) and in general values of $\psi_k(G)$ are exactly cardinalities of minimum vertex covers of k-uniform 'path hypergraph' H built from G (every path P on k vertices in G gives rise to a hyperedge in H containing vertices of P, see [3]). The task of computing $\psi_k(G)$ is NP-complete for a general graph G and any $k \geq 2$ (one can use a reduction from minimum vertex cover problem given in [3], Theorem 1). However, there do exist classes of graphs for which polynomial-time algorithms are available for all k - for instance trees as shown by Brešar et al. in [3] or graphs with bounded tree-width (see [3], [6]). Apart from that, exact formulae to compute $\psi_k(G)$ are known only for some special classes of graphs and particular values of k. These include for example grids and $k = 3$ (see [4]), rooted product graphs and $k = 3$ (see [10]). In the case of other graph products some bounds and exact values (for $k = 2, 3$) are also given in [11]. The paper [10] gives one of the very few known relations between the numbers $\psi_m(G)$ for different m (namely between the value of $\psi_k(G \circ H)$ in a product graph and values of $\psi_k(G)$, $\psi_k(H)$ and $\psi_2(H)$ of its factors) – in Section 3. we present some new results of this kind.

As for the approximation of path numbers, following [3] one can adapt standard 2-approximation algorithm for the minimum vertex cover problem to approximate $\psi_k(G)$ within a factor of k – but polynomial-time approximation with better constant factor is not known in general for $\psi_k(G)$ and $k > 3$. For $k = 3$ however, there exist algorithms of Tu and Zhou [17], [18] with an approximation factor of 2. Some recent results for approximating $\psi_3(G)$ and $\psi_4(G)$ in cubic graphs are also given in [13], [16]. On the other hand, the famous result of Dinur and Safra [8] shows that it is NP-hard to approximate the value of $\psi_2(G)$ to within any factor smaller than $10\sqrt{5}-21 = 1.3606...$ (unless P = NP). In [7] it is shown that approximation of minimum vertex cover in k-uniform hypergraphs for $k \geq 3$ is NP-hard within a constant factor below $k-1$ and thus it seems unlikely that better approximation of $\psi_k(G)$ can be achieved. Bearing that in mind, it appears difficult to obtain a constant approximation factor for a path sequence as a whole (i.e. approximate all $\psi_k(G)$ within the same factor, independent of k and number of vertices in G), but the problem is worth further research.

2. Elementary results

Unless otherwise stated, in the following G will denote a (simple, nonempty) graph on n vertices and k a positive integer satisfying $1 \leq k \leq n$.

For the standard notations and definitions in graph theory we refer the reader to [2]; here we only recall some extensively used notations. If $G = (V, E)$ is a graph and $S \subset V$, then $G[S]$ denotes the subgraph induced by S. Now for $v \in V$ and $e \in E$ we denote by $G - v$ the graph $G[V \setminus \{v\}]$ and by $G - e$ the graph $(V, E \setminus \{e\})$. We will also write $|G|$ for the number of vertices in G. By P_n, C_n and K_n we denote a path, a cycle and a complete graph on n vertices respectively. A complete bipartite graph with partitions of size a and b will be denoted by $K_{a,b}$. The symbol $\simeq$ denotes graph isomorphism and by 'disjoint graphs' we mean vertex disjoint graphs. For a vertex v of G we denote by $d(v)$ the degree of v and by $N(v)$ the neighbourhood of v (the set of all vertices adjacent to v).

Let us first note that from the definition of path numbers one immediately gets $\psi_1(G) = n$, $\psi_k(G) \leq n - k + 1$ (an arbitrary subset of $n - k + 1$ vertices is a k-PVC) and that a path sequence is non-increasing, i.e. $\psi_1(G) \geq \psi_2(G) \geq \psi_3(G) \geq \ldots \geq \psi_n(G) \geq 0$. An easy calculation gives path numbers for paths, cycles and complete graphs, namely $\psi_k(P_n) = \lfloor \frac{n}{k} \rfloor$, $\psi_k(C_n) = \lceil \frac{n}{k} \rceil$ and $\psi_k(K_n) = n - k + 1$ (see [4]). We present values for complete bipartite graphs below.

Proposition 2 *Let $1 \leq k \leq a + b$. Then:*

$$\psi_k(K_{a,b}) = \begin{cases} a + b & \text{if } k = 1, \\ \min\{a, b\} - \lfloor \frac{k}{2} \rfloor + 1 & \text{for } 1 < k \leq 2\min\{a, b\} + 1, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Let us write A and B for partitions of $K_{a,b}$ with $|A| = a \leq b = |B|$. The case $k = 1$ is clear. Assume that $1 < k \leq a + b$ and take $p = a - \lfloor \frac{k}{2} \rfloor + 1$. Since any path in $K_{a,b}$ alternates between A and B, a path on k vertices must have at least $\lfloor \frac{k}{2} \rfloor$ vertices in common with A. It follows that for $k > 2a + 1$ there is no path on k vertices in $K_{a,b}$ and so $\psi_k(K_{a,b}) = 0$. So let $k \leq 2a + 1$ and note that from the above reasoning an arbitrary set of p vertices in A is a k-PVC and consequently $\psi_k(K_{a,b}) \leq p$.

Now let T be a subset of $A \cup B$ and $|T| = p - 1$. To show that $\psi_k(K_{a,b}) > p - 1$ we will build a path on k vertices disjoint from T. It suffices to find an arbitrary set of $\lfloor \frac{k}{2} \rfloor$ vertices in $A \setminus T$ and $\lceil \frac{k}{2} \rceil$ vertices in $B \setminus T$ (or vice versa). Note that there are at least $a - p + 1 = \lfloor \frac{k}{2} \rfloor$ vertices in $A \setminus T$ and also at least $b - p + 1 \geq \lfloor \frac{k}{2} \rfloor$ vertices in $B \setminus T$, since $b \geq a$. If $b > a$, then $b - p + 1 \geq \lfloor \frac{k}{2} \rfloor + 1 \geq \lceil \frac{k}{2} \rceil$ and the result follows. So we can assume that $a = b$. If $|A \setminus T| > \lfloor \frac{k}{2} \rfloor$, then $|A \setminus T| \geq \lceil \frac{k}{2} \rceil$ and we are done since $|B \setminus T| \geq \lfloor \frac{k}{2} \rfloor$. By symmetry, the only case left is $|A| = |B|$, $|A \setminus T| = |B \setminus T| = \lfloor \frac{k}{2} \rfloor$. But if $|A \setminus T| = \lfloor \frac{k}{2} \rfloor$, then $T \subset A$ and $B \cap T = \emptyset$, so $|B \setminus T| = |B|$. It follows that $|A \setminus T| = |B \setminus T| = |B| = |A|$ and so $T = \emptyset$; consequently $p = 1$ and $a = \lfloor \frac{k}{2} \rfloor$, so $k = 2a$ or $k = 2a + 1$. It is easily verified that $\psi_{2a}(K_{a,a}) = 1$ and the value agrees with the formula given in the proposition; the case $k = 2a + 1$ is impossible since $k \leq a + b = 2a$. $\qquad\square$

Let us also note the following useful lemma:

Lemma 3 *Let $G = (V, E)$ be a graph on $n \geq 2$ vertices, $k < n$ and $v \in V$. Then $\psi_k(G) \leq \psi_k(G - v) + 1$. Moreover, the following conditions are equivalent:*

1. $\psi_k(G) = \psi_k(G - v) + 1$,

2. $\exists S \subset V : S$ *is a minimum k-PVC for G and $v \in S$,*

3. $\exists T \subset V \setminus \{v\} : T$ *is a k-PVC for $G - v$ and $T \cup \{v\}$ is a minimum k-PVC for G.*

Proof. Let U be a minimum k-PVC for $G - v$. Then clearly $U \cup \{v\}$ is a k-PVC for G and so $\psi_k(G) \leq \psi_k(G - v) + 1$.
Now suppose that (1) holds and U is as above - then $U \cup \{v\}$ is a minimum k-PVC for G and (2) follows with $S = U \cup \{v\}$. If (2) holds, then $T = S \setminus \{v\}$ is a k-PVC for $G - v$ (a path disjoint from T in $G - v$ is disjoint from S in G) and (3) clearly follows. Supposing that (3) holds gives $\psi_k(G) = |T| + 1 \geq \psi_k(G - v) + 1 \geq \psi_k(G)$ by the first part of the lemma, so $\psi_k(G) = \psi_k(G - v) + 1$. $\qquad\square$

As a corollary we get:

Corollary 4 *Let $G = (V, E)$ be a graph and $e \in E$. Then $\psi_k(G) \leq \psi_k(G - e) + 1$.*

Proof. Let $e = uv$. Since $G - u$ is a subgraph of $G - e$, we apply previous lemma to obtain $\psi_k(G) \leq \psi_k(G - u) + 1 \leq \psi_k(G - e) + 1$. $\qquad\square$

The following remark shows that there are no restrictions on the structure of a minimum k-PVC.

Remark 5 *For any graph $H = (W, F)$, there exists a supergraph G of H such that $G[W] \simeq H$ and that W is a minimum k-PVC for G.*

Proof. We adapt the construction from the proof of [3], Theorem 1. So let us replace each vertex v of H with a path $P^{(v)} = v - w_2^{(v)} - \ldots - w_k^{(v)}$ on k vertices (paths for different v are pairwise disjoint), leaving edges F intact. Call the resulting graph G and note that $G[W] \simeq H$. Now any k-PVC for G must contain at least one vertex from each path $P^{(v)}$, so $\psi_k(G) \geq |W|$. But W is clearly a k-PVC, so $\psi_k(G) = |W|$. $\qquad\square$

3. Two element subsequences

In this section we investigate relations between two path numbers $\psi_k(G)$ and $\psi_m(G)$ for an arbitrary graph G. Let us start with the following example, showing that two elements of a path sequence must satisfy some additional conditions apart from the ones given in the previous section.

Example 6 *There is no graph G satisfying $\psi_{10}(G) = 2$ and $\psi_2(G) = 5$.*

Proof. Suppose that such a graph G exists. Let S be a minimum 2-PVC for G and $v \in S$. Note that at least one vertex from any edge in G belongs to S. Take an arbitrary path P in G that avoids v. Since for every two consecutive vertices on P at least one is from S and P has no more than 4 vertices in common with S – it follows that P is a path on at most 9 vertices. This shows that $\{v\}$ is a 10-PVC for G and so $\psi_{10}(G) \leq 1$, a contradiction. $\square$

These additional necessary conditions are presented in the following proposition.

Proposition 7 *Let $1 \leq m < k$ and $\psi_k(G) > 0$. Then $\psi_m(G) \geq \psi_k(G) + \lfloor \frac{k}{m} \rfloor - 1$.*

Proof. We will proceed by induction on n - the number of vertices in G. The result clearly follows for $n \leq 2$ and also for all graphs G with $\psi_k(G) = 1$ (including the case $k = n$), because we have $\psi_m(G) \geq \psi_m(P_k) = \lfloor \frac{k}{m} \rfloor$. So we can assume that $n > k \geq 2$ and $\psi_k(G) > 1$. Let S be a minimum m-PVC for G and $v \in S$. By Lemma 3, we get $\psi_m(G) = \psi_m(G-v)+1$ and $\psi_k(G-v) \geq \psi_k(G)-1 > 0$. By the induction hypothesis $\psi_m(G - v) \geq \psi_k(G - v) + \lfloor \frac{k}{m} \rfloor - 1$ and consequently $\psi_m(G) = \psi_m(G - v) + 1 \geq \psi_k(G - v) + \lfloor \frac{k}{m} \rfloor \geq \psi_k(G) - 1 + \lfloor \frac{k}{m} \rfloor$. $\square$

Remark 8 *Notice that the condition $\psi_k(G) > 0$ in the above proposition cannot be omitted: take for example $G = K_{1,8}$ (a star on 9 vertices), $k = 9$ and $m \in \{2, 3, 4\}$.*

Remark 9 *Let $m < k$ and take G equal to s disjoint copies of P_m. Then clearly $\psi_m(G) = s$ and $\psi_k(G) = 0$. This shows that there exist graphs G with $\psi_k(G) = 0$ and an arbitrary value of $\psi_m(G)$.*

As the converse of Proposition 7 we show the following:

Theorem 10 *Let k be a positive integer and $1 \leq m < k$. If two integers p_k, p_m satisfy: $p_k > 0$ and $p_m \geq p_k + \lfloor \frac{k}{m} \rfloor - 1$, then there exists a (connected) graph G such that $\psi_k(G) = p_k$ and $\psi_m(G) = p_m$.*

Proof. Let $\lfloor \frac{k}{m} \rfloor = a$, i.e. $am \leq k < (a + 1)m$. Take $p_k + a - 1$ disjoint paths $P^{(1)}, P^{(2)}, \ldots, P^{(p_k+a-1)}$ on $2m-1$ vertices and let $P^{(i)} = v_1^{(i)} - v_2^{(i)} - \ldots - v_{2m-1}^{(i)}$. Now add edges connecting vertices $v_x^{(i)}$ and $v_m^{(j)}$ for all $i \neq j$ and all x, i.e. $1 \leq x \leq 2m-1$. Call the resulting graph H and let $M = \{v_m^{(i)} : 1 \leq i \leq p_k + a - 1\}$ denotes the set of "middle" vertices of all $P^{(i)}$ (see Fig. 1). Since $\psi_m(P^{(i)}) = 1$, it is easily seen that M is a minimum m-PVC for H and so $\psi_m(H) = |M| = p_k + a - 1$.

Let us now prove the following lemma concerning H and M:

Lemma 11 *Any path on k vertices in H must contain at least a vertices from M.*

Proof. Suppose to the contrary that $W = w_1 - \ldots - w_k$ is a path in H and $|\{w_1, \ldots, w_k\} \cap M| = t < a$. Let us divide W into consecutive fragments (subpaths) contained in paths $P^{(i)}$, i.e. $W = W^1 - \ldots - W^s$, where each W^j is a (maximal) subpath of W with all vertices in some fixed $P^{(i)}$. Note that it is possible for multiple W^j to be subpaths of a single $P^{(i)}$ and that any W^j with at least m vertices must contain some vertex from M. Let us now define two sets of indices j: $I_0 = \{j : W^j \cap M = \emptyset\}$ and $I_1 = \{j : W^j \cap M \neq \emptyset\} = \{j : W^j$ has exactly one vertex in common with $M\}$. Clearly, the total number of W^j equals $s = |I_0| + |I_1| = |I_0| + t$. Now we write I_1

Figure 1. An example graph H for $m = 3$ and $p_k + a - 1 = 3$. Paths $P^{(i)}$ are drawn horizontal, set M is marked in black.
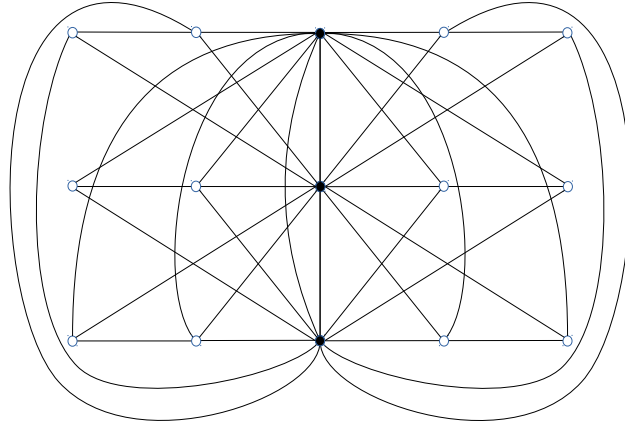
as a disjoint sum of the following subsets: $B = \{j : W^j$ has more than m vertices$\}$, $U = \{j : W^j$ is a single vertex from $M\}$ and $R = I_1 \setminus (B \cup U) = \{j \in I_1 : 2 \leq |W^j| \leq m\}$. Counting the number of vertices in W as a sum of the numbers of vertices in W^j yields the following bound:

$$|W| = \sum_{j \in I_0 \cup I_1} |W^j| = \sum_{j \in I_0 \cup B \cup U \cup R} |W^j| \leq$$
$$\leq |I_0| \cdot (m - 1) + |B| \cdot (2m - 1) + |U| \cdot 1 + |R| \cdot m =$$
$$= |I_0| \cdot (m - 1) + |B| \cdot (2m - 1) + |U| + (s - |I_0| - |B| - |U|) \cdot m =$$
$$= (s + |B| - |U|) \cdot m + |U| - |I_0| - |B|$$

We will now show the following claim: let $a < b$ be two integers such that the last (in order imposed by W) vertex in W^a and the first vertex in W^b are not in M. Then there exists an index $p \in U$ such that $a < p < b$. Indeed, by the construction of H and W^j, since the last vertex in W^a is not in M, the first one in W^{a+1} must be in M. Analogously, the last vertex in W^{b-1} must be in M. If $a + 1 \in U$ or $b - 1 \in U$, then we are done. If not, the last vertex in W^{a+1} and the first one in W^{b-1} are not in M and we can proceed by induction on $b - a$ (the case $b - a = 1$ being impossible and $b - a = 2$ easily verified).

Since any $a, b \in I_0 \cup B$ such that $a < b$ satisfy the hypothesis of the claim, we get $|U| \geq |I_0| + |B| - 1$. The bound for $|W|$ attains its maximum for the smallest possible $|U|$, that is $|U| = |I_0| + |B| - 1$ and then we get $|W| \leq (s - |I_0| + 1) \cdot m - 1 = (t + 1) \cdot m - 1 \leq am - 1 < k$, a contradiction that ends the proof. $\square$

We will show that $\psi_k(H) = p_k$. First note that by Lemma 11 we get $\psi_k(H) \leq p_k$, since the set $S = \{v_m^{(i)} : 1 \leq i \leq p_k\}$ is a k-PVC for H as there are only $a - 1$ vertices in $M \setminus S$. Now let T be an arbitrary set of no more than $p_k - 1$ vertices from H. Without loss of generality we can assume that T has no vertices in common

with paths $P^{(1)}, \ldots, P^{(a)}$ (recall that the number of $P^{(i)}$ is $p_k + a - 1$). It is easy to observe that joining the paths $v_1^{(1)} - \ldots - v_{2m-1}^{(1)}$, $v_m^{(2)} - v_{m+1}^{(2)} - \ldots - v_{2m-1}^{(2)}, \ldots, v_m^{(a)} - v_{m+1}^{(a)} - \ldots - v_{2m-1}^{(a)}$ results in a path on $2m - 1 + (a - 1)m = (a + 1)m - 1 \geq k$ vertices. Consequently T is not a k-PVC for H and $\psi_k(H) > p_k - 1$, so we must have $\psi_k(H) = p_k$.

Now we deal with the m-path number. Take $p_m - (p_k + a - 1)$ (by assumption this number is non-negative) disjoint paths $Q^{(1)}, Q^{(2)}, \ldots, Q^{(p_m - (p_k + a - 1))}$ on m vertices and let $Q^{(j)} = u_1^{(j)} - \ldots - u_m^{(j)}$. Connect all $Q^{(j)}$ to the vertex $v_m^{(1)}$ of $P^{(1)}$ by adding edges $u_1^{(j)} - v_m^{(1)}$ for all j. Resulting graph G satisfies $\psi_k(G) = \psi_k(H) = p_k$, since $S = \{v_m^{(i)} : 1 \leq i \leq p_k\}$ is a k-PVC for G. But also $\psi_m(G) = \psi_m(H) + (p_m - (p_k + a - 1)) = p_m$ because at least one vertex from each $P^{(i)}$ and each $Q^{(j)}$ must be included in the minimum m-PVC of G and clearly $M \cup \{u_1^{(j)} : j = 1, 2, \ldots, p_m - (p_k + a - 1)\}$ is a m-PVC for G. $\square$

4. Path sequences for small graphs

In this section we give some properties of path sequences concerning graphs with small number of vertices. First problem which arises naturally is the question whether equality of path sequences implies graph isomorphism. This is true for graphs with at most three vertices but false in general, as shown by the proposition below.

Proposition 12 *For any $n \geq 4$ there exist (connected) graphs G, H on n vertices such that $\psi(G) = \psi(H)$ but G and H are not isomorphic.*

Proof. Let G be a graph and v a vertex in G. By $G_{v,k}$ we understand a graph obtained from G by adding k new vertices $\{u_1, \ldots, u_k\}$ and edges $\{vu_1, vu_2, \ldots, vu_k\}$ to G. Now, let $u \in V(C_4)$ and let $v \in V(K_4 - e)$ be of degree 3 and consider graphs $G = (C_4)_{u,n-4}$ and $H = (K_4 - e)_{v,n-4}$. Obviously G and H are not isomorphic but

$$\psi(G) = \psi(H) = \begin{cases} (4, 2, 2, 1) & \text{if } n = 4, \\ (n, 2, 2, 1, 1, 0, \ldots, 0) & \text{for } n \geq 5. \end{cases}$$

$\square$

Before going further we state the following definition:

Definition 13 *Let $(p_1, \ldots, p_n)$ be a sequence of non-negative integers. Put*

$$\mathrm{m}(p_1, \ldots, p_n) := \quad \begin{aligned} &\textit{number of non-isomorphic connected graphs } G \\ &\textit{on } n \textit{ vertices such that } \psi(G) = (p_1, \ldots, p_n). \end{aligned}$$

We will call this number the path multiplicity of a sequence $(p_1, \ldots, p_n)$. A sequence with nonzero path multiplicity will be called realisable, i.e. $(p_1, \ldots, p_n)$ is realisable if there exists a connected graph G with $\psi(G) = (p_1, \ldots, p_n)$. Moreover if at least one of the graphs G satisfying $\psi(G) = (p_1, \ldots, p_n)$ is a tree, a bipartite graph, etc. we will say that the sequence is realisable by a tree, a bipartite graph, etc.

Tables 1 and 2 give realisable sequences and their path multiplicities for connected graphs on $n = 5, 6$ and 7 vertices (for smaller n all values are easily found 'by hand'). These numbers were generated using a computer program written by the authors – source code is available at [20]. The lists of non-isomorphic connected graphs and trees were obtained by the Mathematica package [19] and data from the web page [14]. Note: sequences realisable by trees are marked with *.

Table 1. Path sequences for all connected graphs on 5 vertices (9 sequences, 21 graphs) and 6 vertices (20 sequences, 112 graphs).

Sequence	Multiplicity
*(5,1,1,0,0)	1
*(5,2,1,1,0)	2
*(5,2,1,1,1)	1
(5,2,2,1,1)	5
(5,3,1,1,1)	2
(5,3,2,1,1)	2
(5,3,2,2,1)	5
(5,3,3,2,1)	2
(5,4,3,2,1)	1

Sequence	Mult.	Sequence	Mult.
*(6,1,1,0,0,0)	1	(6,3,3,2,1,1)	5
*(6,2,1,1,0,0)	2	(6,3,3,2,2,1)	14
*(6,2,2,1,0,0)	1	(6,4,2,1,1,1)	4
*(6,2,2,1,1,0)	10	(6,4,2,2,1,1)	1
*(6,3,1,1,1,0)	3	(6,4,2,2,2,1)	8
(6,3,2,1,1,0)	3	(6,4,3,2,1,1)	2
*(6,3,2,1,1,1)	9	(6,4,3,2,2,1)	7
(6,3,2,2,1,0)	1	(6,4,3,3,2,1)	9
(6,3,2,2,1,1)	22	(6,4,4,3,2,1)	3
(6,3,2,2,2,1)	6	(6,5,4,3,2,1)	1

Table 2. Path sequences for all connected graphs on 7 vertices (50 sequences, 853 graphs).

| Sequence	Multiplicity	Sequence	Mult.	Sequence	Mult.
*(7,1,1,0,0,0,0)	1	(7,3,3,2,2,1,0)	1	(7,4,4,3,2,1,1)	6
*(7,2,1,1,0,0,0)	2	(7,3,3,2,2,1,1)	87	(7,4,4,3,2,2,1)	24
*(7,2,2,1,0,0,0)	1	(7,4,1,1,1,0,0)	3	(7,4,4,3,3,2,1)	36
*(7,2,2,1,1,0,0)	16	(7,4,2,1,1,1,0)	6	(7,5,3,1,1,1,1)	3
*(7,3,1,1,1,0,0)	4	(7,4,2,1,1,1,1)	3	(7,5,3,2,1,1,1)	4
*(7,3,2,1,1,0,0)	5	(7,4,2,2,1,1,0)	1	(7,5,3,2,2,1,1)	1
*(7,3,2,1,1,1,0)	21	(7,4,2,2,1,1,1)	25	(7,5,3,2,2,2,1)	18
*(7,3,2,1,1,1,1)	2	(7,4,2,2,2,1,1)	39	(7,5,3,3,2,1,1)	1
(7,3,2,2,1,0,0)	1	(7,4,3,1,1,1,0)	1	(7,5,3,3,2,2,1)	8
(7,3,2,2,1,1,0)	39	(7,4,3,1,1,1,1)	12	(7,5,3,3,3,2,1)	22
(7,3,2,2,1,1,1)	4	(7,4,3,2,1,1,0)	3	(7,5,4,3,2,1,1)	2
(7,3,2,2,2,1,0)	1	(7,4,3,2,1,1,1)	20	(7,5,4,3,2,2,1)	7
(7,3,2,2,2,1,1)	9	(7,4,3,2,2,1,1)	69	(7,5,4,3,3,2,1)	19
(7,3,3,1,1,1,0)	3	(7,4,3,2,2,2,1)	81	(7,5,4,4,3,2,1)	15
(7,3,3,1,1,1,1)	8	(7,4,3,3,2,1,1)	46	(7,5,5,4,3,2,1)	3
(7,3,3,2,1,1,0)	10	(7,4,3,3,2,2,1)	129	(7,6,5,4,3,2,1)	1
(7,3,3,2,1,1,1)	10	(7,4,3,3,3,2,1)	20	-	-

From Tables 1 and 2 we can draw some observations. First of all notice that from the basic properties of path numbers it follows that $m(n, n-1, n-2, \ldots, 1) = 1$ (realisable by K_n) and that $m(n, 1, 1, 0, \ldots, 0) = 1$ (realisable by $K_{1,(n-1)}$, for $n \geq 3$). However there are many other path sequences with multiplicity one, for example $m(5, 2, 1, 1, 1) = 1$.

Proposition 12 shows that equality of path sequences for two graphs does not imply that they are isomorphic. The tables above show that we can even have that $\psi(G) = \psi(T)$ for some tree T and some non-tree graph G.

However if T_1, T_2 are trees with $n < 7$ vertices, then it follows from Tables 1 and 2 (and the number of non-isomorphic trees on n vertices) that $\psi(T_1) = \psi(T_2) \iff T_1 \simeq T_2$. But this is also not true in general, as the following proposition shows.

Proposition 14 *For any $n \geq 7$ there exist trees T_1, T_2 on n vertices such that $\psi(T_1) = \psi(T_2)$ but T_1 and T_2 are not isomorphic.*

Proof. Assume first that $n = 7$ and take the following trees:
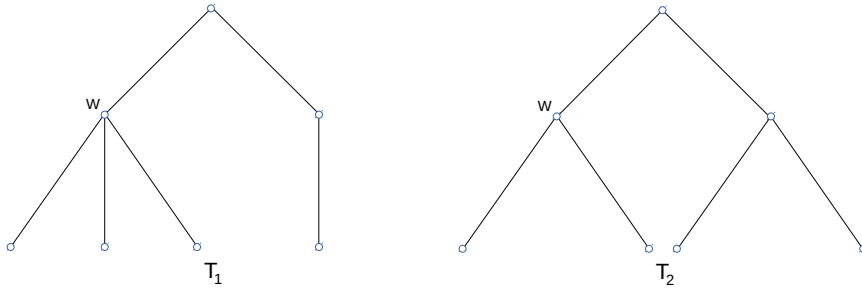
that are clearly not isomorphic and have path sequences equal to $(7, 2, 2, 1, 1, 0, 0)$. For $n > 7$ it suffices to attach additional vertices to w (consider $(T_1)_{w,n-7}$ and $(T_2)_{w,n-7}$ in notations of Proposition 12) to obtain non-isomorphic trees with path sequences $(n, 2, 2, 1, 1, 0, \ldots, 0)$. $\qquad\square$

By analysing the data for graphs with up to 9 vertices (path sequences were calculated with the help of the computer program [20]) we state the following conjecture concerning the existence of a Hamilton path in G (that is clearly equivalent to the condition $\psi_n(G) = 1$). According to our knowledge this conjecture has not been studied yet.

Conjecture 15 *Let G be a connected graph on $n \geq 2$ vertices. Then the following implication holds*

$$\psi_{n-1}(G) = 2 \Rightarrow \psi_n(G) = 1$$

By setting $k = n-1$ in the following remark, one observes that it is not necessary to formulate Conjecture 15 for disconnected graphs.

Remark 16 *If G is a graph on n vertices such that $\psi_k(G) = n - k + 1$ for some $2 \leq k \leq n$, then G is connected.*

Proof. Let $n \geq 2$ and let H_i, $i \in I$ be the connected components of G. Since $\psi_k(H_i) \leq |H_i| - k + 1$, we get $\psi_k(G) = \sum_{i \in I} \psi_k(H_i) \leq \sum_{i \in I}(|H_i| - k + 1) = n - (k - 1) \cdot |I|$. Now if $\psi_k(G) = n - k + 1$, then $n - k + 1 \leq n - (k - 1) \cdot |I|$. Since $k \geq 2$, equality is possible only for $|I| = 1$, i.e. when G is connected. $\qquad\square$

It is straightforward to see that if G is a graph on n vertices and $\psi_2(G)$ is maximum possible (i.e. $n-1$), then G is necessarily isomorphic to K_n. It is not the case for

$\psi_k(G)$ and $k > 2$, however Conjecture 15 implies the following interesting property of path sequences.

Theorem 17 *Let G be a graph on $n \geq 3$ vertices and $2 \leq k < n$.*

If Conjecture 15 holds for all connected graphs with at most n vertices, then $\psi_k(G) = n - k + 1$ implies $\psi_j(G) = n - j + 1$ for all j such that $k < j \leq n$.

Proof. It is enough to prove the following claim for all graphs G on $n \geq 3$ vertices and all m such that $2 \leq m < n$: if Conjecture 15 holds for all connected graphs with at most n vertices, then $\psi_m(G) = n - m + 1$ implies $\psi_{m+1}(G) = n - m$.

We proceed by induction on n. It easy to check the claim for $n = 3$. So assume that $n \geq 4$ and fix $m \in \{2, 3, \ldots, n - 1\}$. Let G be a graph on n vertices such that $\psi_m(G) = n - m + 1$ and $\psi_{m+1}(G) = n - m - t$, with some $t \geq 0$. We need to prove that $t = 0$. Observe that G is connected by Remark 16 and if $m = n - 1$, then the result follows by the validity of Conjecture 15. So assume that $m < n - 1$ and put S to be a minimum $(m + 1)$-PVC for G. There are two cases to consider:

(a) $S \neq \emptyset$. Choose $v \in S$ and let $G' = G - v$. Lemma 3 gives $n - m \leq \psi_m(G') \leq (n - 1) - m + 1$ and therefore $\psi_m(G') = n - m$. By the induction hypothesis $\psi_{m+1}(G') = n - m - 1$, but due to Lemma 3 we obtain that $\psi_{m+1}(G') = \psi_{m+1}(G) - 1 = n - m - t - 1$, so $t = 0$.

(b) $S = \emptyset$. This case cannot occur and the proof is as follows: let w be any vertex of G, put $G' = G - w$ and observe that by Lemma 3 we get $\psi_m(G') = n - m$, so by the induction hypothesis $\psi_{m+1}(G') = n - m - 1$. But $0 \leq \psi_{m+1}(G') \leq \psi_{m+1}(G) = 0$ and consequently $n - m - 1 = 0$, which is impossible since $m < n - 1$. $\qquad\square$

We now present a lemma which will be useful in giving a direct proof of Conjecture 15 for graphs with no more than 7 vertices. Note that the points (2)–(4) follow from Lemma 2.2 of [5], however we include them here with a proof.

Lemma 18 *Let $G = (V, E)$ be a connected graph on at least four vertices such that $V = \{p_1, \ldots, p_{n-1}, q\}$, $N(q) = \{p_{i_1}, \ldots, p_{i_t}\}$ and $p_1 - \ldots - p_{n-1}$ is a path in G. If $\psi_{n-1}(G) = 2$ and $\psi_n(G) = 0$, then*

(1) $d(p_1) \geq 2$ and $d(p_{n-1}) \geq 2$.

(2) $p_1 p_{n-1} \notin E$, $q p_1 \notin E$, $q p_{n-1} \notin E$ and if $q p_i \in E$, then $q p_{i+1} \notin E$ for all $i \in \{2, \ldots, n - 2\}$.

(3) $p_1 p_{i_j + 1} \notin E$ and $p_{n-1} p_{i_j - 1} \notin E$ for all $j \in \{1, \ldots, t\}$.

(4) $p_1 p_{i_j - 1} \notin E$, for all j such that $i_j > \min\{i_1, \ldots, i_t\}$ and $p_{n-1} p_{i_j + 1} \notin E$ for all j such that $i_j < \max\{i_1, \ldots, i_t\}$.

Proof. To see (1) suppose to the contrary that $d(p_1) = 1$. We show that $S = \{p_2\}$ is a $(n - 1)$-PVC for G. This follows from the fact that any path on k vertices which avoids p_2 must also avoid p_1 and therefore $k \leq n - 2$. The second case is proved analogously.

Now, the first part of (2) is obvious (in any case we get a path on all vertices in G, which contradicts $\psi_n(G) = 0$). For the second, if such an i exists, we have a path $p_1 - \ldots - p_i - q - p_{i+1} - \ldots - p_{n-1}$.

As far as (3) is concerned, suppose first that $p_1 p_{i_j+1} \in E$ for some i_j such that $p_{i_j} \in N(q)$. Then we have the following path on n vertices in G: $q - p_{i_j} - p_{i_j-1} - \ldots - p_1 - p_{i_j+1} - p_{i_j+2} - \ldots - p_{n-1}$. If now $p_{n-1}p_{i_j-1} \in E$, then we get that the following path on n vertices: $p_1 - \ldots - p_{i_j-1} - p_{n-1} - \ldots - p_{i_j} - q$ exists.

To prove (4), let $r = \min\{i_1, \ldots, i_t\}$ and suppose that $p_1 p_{i_j-1} \in E$. Then we have the following path on n vertices in G: $p_{n-1} - p_{n-2} - \ldots - p_{i_j} - q - p_r - p_{r-1} - \ldots - p_1 - p_{i_j-1} - p_{i_j-2} - \ldots - p_{r+1}$. The second case follows by symmetry argument. $\square$

Corollary 19 *Let $G = (V, E)$ be a connected graph on at least four vertices such that $V = \{p_1, \ldots, p_{n-1}, q\}$ and $p_1 - \ldots - p_{n-1}$ is a path in G. If $\psi_{n-1}(G) = 2$ and $\psi_n(G) = 0$, then*

$$2 \le d(q) \le \left\lceil \frac{n-3}{2} \right\rceil$$

Proof. Firstly notice that $d(q) \ge 2$, since if $d(q) = 1$ and $qu \in E$, then $S = \{u\}$ is a $(n-1)$-PVC. To see this note that any path P in G that avoids u must also avoid q and so $|P| < n - 1$. As for the second inequality, it follows from Lemma 18.(2) since for every $i \in \{2, \ldots, n-2\}$ at most one of the edges qp_i, qp_{i+1} is in E. $\square$

The above facts allow us to give a direct proof of Conjecture 15 for graphs with no more than 7 vertices. Our reasoning is 'by considering cases' – unfortunately we were unable to find a more general approach.

Theorem 20 *Let $G = (V, E)$ be a connected graph on n vertices, with $2 \le n \le 7$. Then Conjecture 15 holds for G, i.e.*

$$\psi_{n-1}(G) = 2 \Rightarrow \psi_n(G) = 1$$

Proof. The theorem holds true for $n = 2$ and $n = 3$, with complete graphs K_2 and K_3 being the only cases to verify. So we can assume that $n \ge 4$.

It is sufficient to prove that the existence of a connected graph that satisfies $\psi_{n-1}(G) = 2$ and $\psi_n(G) = 0$ leads to a contradiction. Throughout we consider a graph $G = (V, E)$ with $V = \{p_1, \ldots, p_{n-1}, q\}$ and assume that $p_1 - \ldots - p_{n-1}$ is a path in G.

Notice that if $n = 4$ or $n = 5$, then $\left\lceil \frac{n-3}{2} \right\rceil < 2$ and Corollary 19 gives $2 \le d(q) < 2$, a contradiction. So we can assume that $n = 6$ or $n = 7$. Note that in both cases we get $d(q) = 2$ by Corollary 19.

Now let G be a connected graph with 6 vertices such that $\psi_5(G) = 2$, $\psi_6(G) = 0$ and $d(q) = 2$. Due to Lemma 18 we only need to consider the case when $qp_2, qp_4 \in E$ - but then the set $S_5 = \{p_2\}$ is a 5-PVC. To see this let us assume that there exists a path P on 5 vertices p_1, p_3, p_4, p_5, q (in any order). Using again Lemma 18 we obtain that $d(p_1) = 2, p_1 p_4 \in E$ and $p_3 p_5 \notin E$. It follows that $p_2 p_3$ and $p_3 p_4$ are the only two edges containing p_3. Consequently P must start with p_3 and it is easy to see that we cannot build a path avoiding p_2 longer than $p_3 - p_4 - x$ where $x \in \{q, p_1, p_5\}$, which contradicts P having 5 vertices.

Let us now assume that $n = 7$ and there exists a graph G such that $\psi_6(G) = 2$ and $\psi_7(G) = 0$ with $d(q) = 2$. Because of Lemma 18 and symmetries we only need to consider two cases: $qp_2, qp_5 \in E$ and $qp_2, qp_4 \in E$. Assume the first case – by Lemma 18 we get that $d(p_1) = 2$ and $p_1 p_5 \in E$. But now $S_6 = \{p_2\}$ is a 6-PVC for

G: this follows from the fact that there is no path on 6 vertices which avoids p_2 in G. Indeed, if such a path exists, then it is of the form $q - p_5 - x_1 - x_2 - x_3 - x_4$ where $x_i \neq p_1$ for $i = 1, 2, 3, 4$ – a contradiction.

Let us now proceed with the case $qp_2, qp_4 \in E$. By Lemma 18 we must have that $d(p_1) = 2$ and $p_1p_4 \in E$. But then again $S_6 = \{p_2\}$ is a 6-PVC for G. To see this, notice that we only need to consider paths of the form $q - p_4 - x_1 - x_2 - x_3 - x_4$. If such a path omits p_2, then we cannot have $x_1 = p_1$ and so we get $x_i \in \{p_3, p_5, p_6\}$ for $i = 1, 2, 3, 4$ – a contradiction. $\square$

As a consequence of the above and Theorem 17 we get

Corollary 21 *Let G be a graph on n vertices with $3 \leq n \leq 7$ and let $2 \leq k < n$. If $\psi_k(G) = n - k + 1$, then $\psi_j(G) = n - j + 1$ for all j such that $k < j \leq n$.*

5. References

[1] Acharya H.B., Choi T., Bazzi R.A., Gouda M.G., *The K-Observer Problem in Computer Networks*. In: Stabilization, Safety and Security of Distributed Systems, Lecture Notes in Computer Science, Springer, Berlin-Heidelberg, 2011, 6976, pp. 5–18.

[2] Bollobás B., *Modern Graph Theory*, GTM 184, Springer-Verlag, New York 2002.

[3] Brešar B., Kardoš F., Katrenič J., Semanišin G., *Minimum k-path vertex cover*, Discrete Applied Mathematics, 2011, 159(12), pp. 1189–1195.

[4] Brešar B., Jakovac M., Katrenič J., Semanišin G., Taranenko A., *On the vertex k-path cover*, Discrete Applied Mathematics, 2013, 161(13–14), pp. 1943–1949.

[5] Bullock F., Frick M., Semanišin G., Vlačuha R., *Nontraceable detour graphs*. Discrete Mathematics, 2007, 307(7-8), pp. 839–853.

[6] Courcelle B., *The Monadic Second-Order Logic of Graphs. I. Recognizable Sets of Finite Graphs*. Information and Computation, 1990, 85(1), pp. 12–75.

[7] Dinur I., Guruswami V., Khot S., Regev O., *A New Multilayered PCP and the Hardness of Hypergraph Vertex Cover*. SIAM Journal on Computing, 2005, 34(5), pp. 1129–1146.

[8] Dinur I., Safra S., *On the hardness of approximating minimum vertex cover*. Annals of Mathematics Second Series, 2005, 162(1), pp. 439–485.

[9] Funke S., Nusser A., Storandt S., *On k-Path Covers and their Applications*. Proceedings of the VLDB Endowment, 2014, 7(10), pp. 893–902.

[10] Jakovac M., *The k-path vertex cover of rooted product graphs*. Discrete Applied Mathematics, 2015, 187, pp. 111–119.

[11] Jakovac M., Taranenko A., *On the k-path vertex cover of some graph products.* Discrete Mathematics, 2013, 313(1), pp. 94–100.

[12] Kardoš F., Katrenič J., Schiermeyer I., *On computing the minimum 3-path vertex cover and dissociation number of graphs.* Theoretical Computer Science, 2011, 412(50), pp. 7009–7017.

[13] Li Y., Tu J., *A 2-approximation algorithm for the vertex cover P_4 problem in cubic graphs.* International Journal of Computer Mathematics, 2014, 91(10), pp. 2103–2108.

[14] McKay B., *Home page at the Research School of Computer Science, Australian National University.* Accessed July 22, 2015, https://cs.anu.edu.au/people/Brendan.McKay/data/graphs.html.

[15] Novotný M., *Design and Analysis of a Generalized Canvas Protocol.* In: Information Security Theory and Practices. Security and Privacy of Pervasive Systems and Smart Devices, Lecture Notes in Computer Science vol. 6033, Springer Berlin Heilderberg, 2010, pp. 106–121.

[16] Tu J., Yang F., *The vertex cover P_3 problem in cubic graphs.* Information Processing Letters, 2013, 113(13), pp. 481–485.

[17] Tu J., Zhou W., *A factor 2 approximation algorithm for the vertex cover P_3 problem.* Information Processing Letters, 2011, 111(14), pp. 683–686.

[18] Tu, J., Zhou, W., *A primal-dual approximation algorithm for the vertex cover P_3 problem,* Theoretical Computer Science, 2011, 412(50), pp. 7044–7048.

[19] Wolfram Research, Inc., Mathematica, Version 10.0, Champaign, IL (2014).

[20] Zygadło J., *Home page at the Institute of Computer Science and Computational Mathematics, Jagiellonian University.* http://www.ii.uj.edu.pl/~zygadlo/psik-source.zip.