Discrete Mathematics Letters
www.dmlett.com

Discrete Math. Lett. **7** (2021) 66–73
DOI: 10.47443/dml.2021.0042

Research Article

Vertex identification in trees

Yuya Kono, Ping Zhang*

Department of Mathematics, Western Michigan University, Kalamazoo, Michigan 49008-5248, USA

(Received: 23 April 2021. Accepted: 13 July 2021. Published online: 16 July 2021.)

Abstract

A red-white coloring of a nontrivial connected graph G of diameter d is an assignment of red and white colors to the vertices of G where at least one vertex is colored red. Associated with each vertex v of G is a d-vector, called the code of v, whose ith coordinate is the number of red vertices at distance i from v. A red-white coloring of G for which distinct vertices have distinct codes is called an identification coloring or ID-coloring of G. A graph G possessing an ID-coloring is an ID-graph. The minimum number of red vertices among all ID-colorings of an ID-graph G is the identification number or ID-number of G. Necessary conditions are established for those trees that are ID-graphs. A tree T is starlike if T is obtained by subdividing the edges of a star of order 4 or more. It is shown that for every positive integer r different from 2, there exist starlike trees satisfying some prescribed properties having ID-number r.

Keywords: distance; vertex identification; identification coloring; tree.

2020 Mathematics Subject Classification: 05C05, 05C12, 05C15, 05C90.

1. Introduction

Over the years, many methods have been introduced with the goal of uniquely identifying the vertices of a connected graph. Often these approaches have employed distance and coloring. The oldest of these methods deal with what is referred to as the metric dimension of a connected graph. For a nontrivial connected graph G of order n, the goal is to locate an ordered set $W = \{w_1, w_2, \ldots, w_k\}$ of k vertices in G, $1 \le k \le n$, and associate with each vertex v of G the k-vector $(a_1, a_2, \ldots, a_k)$, where a_i is the distance $d(v, w_i)$ between v and w_i, $1 \le i \le k$. If the n k-vectors produced in this manners are distinct, then the vertices of G have been uniquely identified. For each connected graph G, such a set W can always be found since we can always choose $W = V(G)$. The primary problem here is to determine the minimum size of such a set W. This is referred to as the *metric dimension* of G. Equivalently, the metric dimension of a connected graph G can be defined as the minimum number of vertices of G that can be assigned the same color, say red, such that for every two vertices u and v of G, there exists a red vertex w such that $d(u, w) \ne d(v, w)$. This parameter is defined for every connected graph.

Another method that has been studied to uniquely identify the vertices of a connected graph G has been referred to as the partition dimension of G. For a nontrivial connected graph G of order n, the goal is to obtain a k-coloring, $1 \le k \le n$, of the vertices of G, where the coloring is not required (or expected) to be a proper coloring. This results in k color classes $V_1, V_2, \ldots, V_k$ of $V(G)$. For each vertex v of G, we once again associate a vector, here a k-vector $(a_1, a_2, \ldots, a_k)$ where a_i denotes the distance from v to a nearest vertex in V_i for $1 \le i \le k$. If the vertices of G have distinct k-vectors, then the vertices of G have been uniquely identified. Such a coloring always exists since we can always assign distinct colors to the vertices of G, thereby obtaining a procedure that has similarity to metric dimension. The minimum number of colors that accomplishes this goal is referred to as the *partition dimension of* G. The partition dimension of a connected graph G can also be defined as the minimum number k of colors (denoted by $1, 2, \ldots, k$) that can be assigned to the vertices of G, one color to each vertex, so that for every two vertices u and v of G, there exists a color i such that the distance between u and a nearest vertex colored i is distinct from the distance between v and a nearest vertex colored i. This parameter is also defined for every connected graph.

Another method that has been introduced for the purpose of uniquely identifying the vertices of a connected graph is referred to as an identification coloring. Let G be a connected graph of diameter $d \ge 2$ and let there be given a red-white vertex coloring c of the graph G where at least one vertex is colored red. That is, the *color* $c(v)$ of a vertex v in G is either red or white and $c(v)$ is red for at least one vertex v of G. With each vertex v of G, there is associated a *d-vector* $\vec{d}(v) = (a_1, a_2, \ldots, a_d)$ called the *code* of v corresponding to c, where the ith coordinate a_i is the number of red vertices at

*Corresponding author (ping.zhang@wmich.edu).

distance i from v for $1 \leq i \leq d$. If distinct vertices of G have distinct codes, then c is called an *identification coloring* or *ID-coloring*. Equivalently, an identification coloring of a connected graph G is an assignment of the color red to a nonempty subset of $V(G)$ (with the color white assigned to the remaining vertices of G) such that for every two vertices u and v of G, there is an integer k with $1 \leq k \leq d$ such that the number of red vertices at distance k from u is different from the number of red vertices at distance k from v. A graph possessing an identification coloring is an *ID-graph*. A major difference here from the two methods described above is that not all connected graphs are ID-graphs.

The concept of metric dimension was introduced independently by Slater [15] and by Harary and Melter [10] and has been studied by many (see [4,7], for example). Slater described the usefulness of these ideas when working with U.S. Coast Guard Loran (long range aids to navigation) stations in [15, 16]. Johnson [13, 14] of the former Upjohn Pharmaceutical Company applied this in attempts to develop the capability of large datasets of chemical graphs. The concept of partition dimension was introduced in [6]. These concepts as well as other methods of vertex identifications in graphs have been studied by many with various applications (see [2, 3, 8, 9, 11, 12, 17, 18] for example). The concepts of ID-colorings and ID-graphs were introduced and studied in [5].

All of the methods mentioned above involve constructing a vertex coloring of a connected graph G with the goal of producing a vertex labeling of G (using vectors of the same size as labels) so that distinct vertices of G have the distinct labels. Consequently, the goal of each of these methods is to obtain an *irregular labeling* of G. The general topic of irregularity in graphs is described in [2]. There is the related topic of obtaining a labeling of G by means of colorings where exactly two vertices of G have the same label. These are called *artiregular labelings*, a topic discussed in [1].

We first present five results obtained in [5] on ID-colorings. For an integer $t \geq 2$, the members of a set S of t vertices of a graph G are called t*-tuplets* (*twins* if $t = 2$ and *triplets* if $t = 3$) if either (1) S is an independent set in G and every two vertices in S have the same neighborhood or (2) S is a clique, that is the subgraph $G[S]$ induced by S is complete and every two vertices in S have the same closed neighborhood.

Proposition 1.1. *Let c be an ID-coloring of a connected graph G. If u and v are twins of G, then $c(u) \neq c(v)$. Consequently, if G is an ID-graph, then G is triplet-free.*

Proposition 1.2. *Let c be a red-white coloring of a connected graph G where there is at least one vertex of each color. If x is a red vertex and y is a white vertex, then $\vec{d}(x) \neq \vec{d}(y)$.*

Theorem 1.1. *A nontrivial connected graph G has $\mathrm{ID}(G) = 1$ if and only if G is a path.*

Theorem 1.2. *A connected graph G of diameter 2 is an ID-graph if and only if $G = P_3$.*

Theorem 1.3. *For a positive integer r, there exists a connected graph G with $\mathrm{ID}(G) = r$ if and only if $r \neq 2$.*

The following result describes a property of ID-colorings.

Proposition 1.3. *Let G be a connected graph with an ID-coloring c. If H is a connected subgraph of G such that (i) H contains all red vertices in G and (ii) $d_G(x, y) = d_H(x, y)$ for every two vertices x and y of H, then the restriction of c to H is an ID-coloring of H.*

Proof. Let $\mathrm{diam}(H) = d$ and let c_H be the restriction of c to H. For a vertex v of H, let $\vec{d}_{c_H}(v) = (a'_1, a'_2, \ldots, a'_d)$ and let $\vec{d}_c(v) = (a_1, a_2, \ldots, a_d, \ldots)$. Notice that if $d = \mathrm{diam}(G)$, then $\vec{d}_c(v) = (a_1, a_2, \ldots, a_d)$ while if $d < \mathrm{diam}(G)$, then $a_t = 0$ for each integer t with $d + 1 \leq t \leq \mathrm{diam}(G)$. Since H contains all red vertices in G and $d_G(v, w) = d_H(v, w)$ for every vertex w of G, it follows that $a_i = a'_i$ for $1 \leq i \leq d$ and so the restriction of c to H is an ID-coloring of H. □

Both conditions stated in the hypothesis of Proposition 1.3 for a connected subgraph H of a graph G are needed. For example, consider the ID-graph G in Figure 1. The subgraph H_1 of G does not contain all red vertices of G while the subgraph H_2 is not distance-preserving. For $i = 1, 2$, the restriction of the ID-coloring c of G to the subgraph H_i of G is not an ID-coloring of H_i (since there are twins both of which are colored white).

Here, our emphasis turns to trees that are ID-graphs, namely *ID-trees*. We investigate structural problems of ID-trees, provide necessary conditions for trees to be ID-trees, and establish a realization result on ID-numbers of ID-trees satisfying some prescribed conditions.

2. ID-colorings of trees

The only t-tuplets, $t \geq 2$, in a tree T are end-vertices of T, all with the same neighbor. As we saw, if T contains triplets, then T is not an ID-tree. If T contains twins and possesses an ID-coloring, then the twins must be colored differently in every ID-coloring. We now see that for trees, the concepts of twins and triplets are special cases of something more general.

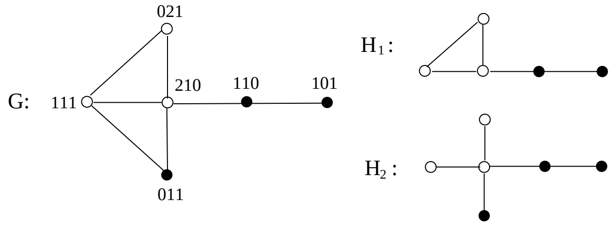

Figure 1: Two subgraphs of an ID-graph G.

If T is a tree with a vertex v possessing two isomorphic branches B_1 and B_2, then B_1 and B_2 are *twin branches* at v if there is an isomorphism from B_1 to B_2 fixing v. If T contains a vertex v possessing three isomorphic branches B_1, B_2, and B_3 such that every two of them are twin branches, then B_1, B_2, and B_3 are *triplet branches* at v. If the size of each branch at v is 1, then T contains twins or triplets. For example, there are three isomorphic branches of size 5 at the vertex v of the tree T of Figure 2. However, T has twin branches at v but no triplet branches at v.

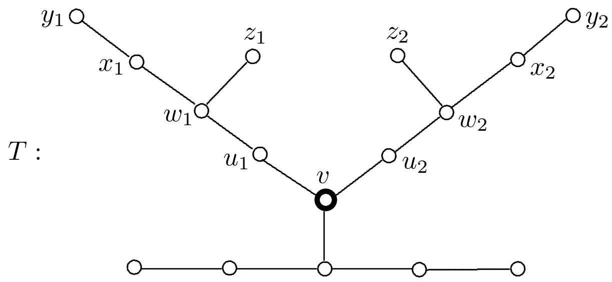

Figure 2: A tree T with twin branches of size 5 at v.

Let T_1 and T_2 be two *rooted* trees whose roots are v_1 and v_2, respectively. Then T_1 and T_2 are considered to be *isomorphic rooted trees*, denoted $T_1 \cong T_2$, if there is an isomorphism $\alpha : V(T_1) \to V(T_2)$ such that $\alpha(v_1) = v_2$. For $i = 1, 2$, let c_i be a red-white coloring of a tree T_i rooted at v_i where $T_1 \cong T_2$. Then c_1 and c_2 are considered to be *isomorphic colorings*, denoted $c_1 \cong c_2$, if there is an isomorphism $\alpha : V(T_1) \to V(T_2)$ such that $\alpha(v_1) = v_2$ and $c_1(x) = c_2(\alpha(x))$ for every vertex x of T_1. In particular, $c_1(v_1) = c_2(v_2)$.

Observation 2.1. *Suppose that a tree T has twin branches B_1 and B_2 at a vertex v and c is a red-white coloring of T. For $i = 1, 2$, let c_i be the restriction of c to B_i rooted at v. If $c_1 \cong c_2$, then c is not an ID-coloring of T.*

Let T_0 be a tree of size $k \geq 1$ rooted at a vertex v. If the color of v is fixed, say v is white, then there are at most 2^k distinct (non-isomorphic) red-white colorings of T_0 *in which v is colored white*. Consequently, if there are more than 2^k copies of a particular branch of size k at v, then T is not an ID-tree by Observation 2.1. In the case when $k = 1$, this simply says that no ID-tree can contain a triplet.

Let T be a tree rooted at a vertex v and let c be an ID-coloring of T. If T_0 is a subtree of T of minimum order rooted at v such that T_0 contains all red vertices in T, then the restriction of c to T_0 is an ID-coloring of T_0 by Proposition 1.3. Necessarily, all end-vertices of T_0 are red. In the case when T_0 is a path P_{k+1} of size k whose end-vertices are v and w, there are at most 2^{k-1} distinct red-white colorings of P_{k+1} in which v is white and w is red.

A tree T is *starlike* if T is obtained by subdividing the edges of a star of order 4 or more. Equivalently, a tree T is starlike if and only if T has exactly one vertex whose degree is greater than 2. This vertex is referred to as the *central vertex* of T. If the degree of the central vertex v of a starlike T is $k \geq 3$, then T has k branches (paths) at v, each branch containing v as an end-vertex of T. For example, the starlike tree T in Figure 3 has four branches at its central vertex. This tree is twin-free but does contain twin branches at its central vertex. This starlike tree is an ID-graph and an ID-coloring having exactly four red vertices is shown in Figure 3. In fact, $\mathrm{ID}(T) = 4$.

Proposition 2.1. *Let T be a starlike tree whose largest branch at its central vertex v has size k. If T is an ID-tree, then for each integer i with $1 \leq i \leq k$, there are at most 2^i branches of size i or less at v. Consequently, if T is an ID-tree, then T has at most 2^k branches at v.*

Proof. In view of Proposition 1.3, it suffices to determine the maximum number of distinct red-white colorings of all branches (paths) of T such that v is white and all end-vertices are red. For each integer i with $1 \leq i \leq k$, there are 2^{i-1} distinct red-white colorings of branches of size i at v in which v is colored white and the other end-vertex of each

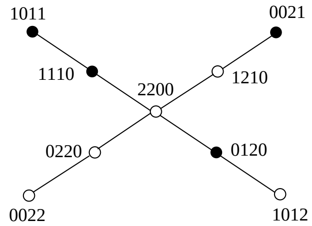

Figure 3: An ID-coloring of a twin-free starlike tree.

branch is colored red. Thus, the minimum number of branches of size i at v without duplicating a red-white coloring of these branches is 2^{i-1}. Therefore, the maximum number of all such red-white colorings of branches of all possibles sizes at v is $\sum_{i=1}^{k} 2^{i-1} = 2^k - 1$. Since there can be one branch of size k or less at v all of whose vertices are colored white, it follows that there can be 2^k branches at v such that the red-white colorings of every two isomorphic branches at v are different. □

Corollary 2.1. *Let T be a starlike tree whose largest branch at its central vertex v has size k. If T has more than 2^k branches at v, then T is not an ID-tree.*

For example, if T is a starlike ID-tree whose largest branch at its central vertex v has size 3, then (1) there are at most two branches of size 1 at v, (2) there are at most four branches of size 2 or less at v, and (3) there are at most eight branches of size 3 or less at v. As an illustration, the three starlike trees of Figure 4 satisfy all conditions (1)–(3).

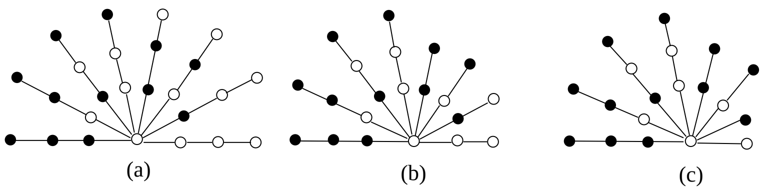

Figure 4: Three starlike trees whose largest branch at its central vertex has size 3.

The tree of Figure 4(a) has eight branches of size 3 at its central vertex and no branches of size less than 3 at its central vertex. The tree of Figure 4(b) has four branches of size 3, four branches of size 2, and no branches of size 1 at its central vertex. The tree of Figure 4(c) has four branches of size 3, two branches of size 2, and two branches of size 1 at its central vertex. In each case, there are eight branches at the central vertex of the tree. The red-white colorings of the three trees in Figure 4 are essentially the same coloring. It can be shown that this coloring is an ID-coloring. For the red-white coloring of the tree T of Figure 4(c), partial codes of the vertices of T containing the initial coordinates of each code are shown in Figure 5. (These partial codes are sufficient to show that all codes are distinct.) Since every two distinct vertices of T have distinct codes, this red-white coloring is an ID-coloring of T.

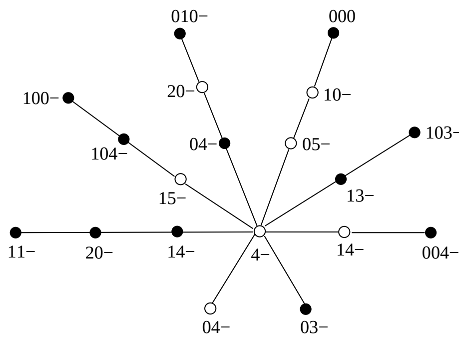

Figure 5: An ID-coloring of a starlike tree.

Theorem 2.1. *If T is a starlike tree with central vertex v whose branches at v have distinct sizes, then T is an ID-tree.*

Proof. Suppose that $\deg v = k \geq 3$ and let $B_1, B_2, \ldots, B_k$ be the branches (paths) of T at v, where B_i has size m_i and $m_i < m_{i+1}$ for $1 \leq i < k$. Define a red-white coloring c of T that assigns the color white to v and the color red to all other vertices of T. We show that c is an ID-coloring of T. By Proposition 1.2, it suffices to show that every two red vertices have distinct codes. Let $x, y \in V(T) - \{v\}$ and let $\vec{d}(x) = (a_1, a_2, \ldots, a_d)$ and $\vec{d}(y) = (b_1, b_1, \ldots, b_d)$ where $d = \text{diam}(T) = m_{k-1} + m_k$. Suppose that $d(x, v) = s$ and $d(y, v) = t$. We consider two cases, according to whether $s \neq t$ or $s = t$.

Case 1. *$s \neq t$, say $s < t$.* Then $a_s \in \{0, 1\}$ and $b_s \in \{1, 2\}$. If $a_s \neq b_s$, then $\vec{d}(x) \neq \vec{d}(y)$. Thus, we may assume that $a_s = b_s = 1$. Thus, $a_{s+1} \in \{k-1, k\}$ and $b_{s+1} \in \{0, 1\}$. Since $k \geq 3$, it follows that $a_{s+1} \geq 2$ and so $a_{s+1} \neq b_{s+1}$, implying that $\vec{d}(x) \neq \vec{d}(y)$.

Case 2. $s = t$. Then x and y belong to different branches of T at v, say $x \in V(B_i)$ and $y \in V(B_j)$ where $1 \leq i < j \leq k$. Let $B_i = (v = v_0, v_1, \ldots, v_{m_i})$ and $B_j = (v = u_0, u_1, \ldots, u_{m_j})$, where then $x = v_s$ and $y = u_s$. If $m_i - s + 1 = s$, then $a_{m_i - s+1} = 0$ and $b_{m_i - s+1} \geq 1$. If $m_i - s + 1 \neq s$, then $b_{m_i - s+1} = a_{m_i - s+1} + 1$. In either case, $\vec{d}(x) \neq \vec{d}(y)$. Therefore, c is an ID-coloring of T. □

3. Starlike trees with prescribed ID-number

We saw in Theorem 1.3 that for every integer $r \geq 3$, there exists a connected graph G with $\text{ID}(G) = r$. For such a given integer r, the graph G described in the proof of Theorem 1.3 contains r pairwise disjoint twins from which it follows that $\text{ID}(G) \geq r$. It was therefore only necessary to show that $\text{ID}(G) \leq r$. We now show that for every integer $r \geq 3$, there exists a tree T with no twins at all such that $\text{ID}(T) = r$. In addition, we show that there is a tree without twin branches having ID-number r. In particular, we show that for every odd integer $r \geq 5$ there is a twin-free tree T whose automorphism group contains $(r+1)!$ elements such that $\text{ID}(T) = r$. We also show that there is a red-white coloring c of the same class of trees T where exactly $r - 1$ vertices are colored red such that $\vec{d}(x) = \vec{d}(y)$ for exactly one pair x, y of vertices of T. Consequently, there is a red-white coloring of these trees T with exactly two vertices having the same code. As we metioned earlier, such a (red-white) coloring results in an *antiregular labeling* (see [1, 2], for example.)

For each integer $r \geq 3$, let $T = S_{r-1}(K_{1,r+1})$ be the starlike tree obtained from the star $K_{1,r+1}$ of order $r+2$ by subdividing each edge of the $r + 1$ edges in K_{r+1} exactly $r - 1$ times. Let v be the central vertex of T. Then the degree of v is $r + 1$ and each of the $r + 1$ branches of T at v has length r. For each integer i with $0 \leq i \leq r$, let $B_i = (v, v_{i,1}, v_{i,2}, \ldots, v_{i,r})$ be a branch of T at v. Then $\text{diam}(T) = 2r$ and T is twin-free.

Theorem 3.1. *For each odd integer $r \geq 3$, $\text{ID}(S_{r-1}(K_{1,r+1})) = r$.*

Proof. For an odd integer $r \geq 3$, let $T = S_{r-1}(K_{1,r+1})$, where v is the central vertex of T and $B_i = (v, v_{i,1}, v_{i,2}, \ldots, v_{i,r})$ is a branch of T at v for $0 \leq i \leq r$. First, we show that $\text{ID}(T) \geq r$. For any red-white coloring of T that assigns the color red to at most $r - 1$ vertices of T, there are at least two branches, say B_0 and B_1, of T at v such that the paths $B_0 - v$ and $B_1 - v$ contain no red vertices of T. However then, $\vec{d}(v_{0,1}) = \vec{d}(v_{1,1})$, for example, and so this red-white coloring is not an ID-coloring of T. Therefore, $\text{ID}(T) \geq r$.

Next, we show that T has an ID-coloring with exactly r red vertices. Define a red-white coloring c of T by assigning the color red to each vertex $v_{i,i}$ for $1 \leq i \leq r$ and white to the remaining vertices of T. Thus, T has exactly r red vertices. It remains to show that c is an ID-coloring of T. Since $\text{diam}(T) = 2r$, the code of each vertex of T is a $(2r)$-vector. Let x and y be two distinct vertices of T. We consider two cases, according to whether x and y are both red or both white. Let $\vec{d}(x) = (a_1, a_2, \ldots, a_{2r})$ and $\vec{d}(y) = (b_1, b_2, \ldots, b_{2r})$.

Case 1. *x and y are both red.* Let $x = v_{i,i}$ and $y = v_{j,j}$ where $1 \leq i < j \leq r$.

- ⋆ First, suppose that $j \neq r$. Since (1) the *last* nonzero coordinate in $\vec{d}(v_{i,i})$ is the $(i + r)$th coordinate where $i + r = d(v_{i,i}, v_{r,r})$ and the *last* nonzero coordinate in $\vec{d}(v_{j,j})$ is the $(j + r)$th coordinate where $j + r = d(v_{j,j}, v_{r,r})$ and (2) $i < j$, it follows that $a_{j+r} = 0$ and $b_{j+r} = 1$ and so $\vec{d}(x) \neq \vec{d}(y)$.

- ⋆ Next, suppose that $j = r$. We saw that the last nonzero coordinate in $\vec{d}(v_{i,i})$ where $1 \leq i \leq r - 1$ is the $(i + r)$th coordinate. Since the last nonzero coordinate in $\vec{d}(v_{r,r})$ is the $(2r - 1)$th coordinate where $2r - 1 = d(v_{r-1,r-1}, v_{r,r})$, it follows that if $i \neq r - 1$, then $\vec{d}(x) \neq \vec{d}(y)$. Thus, we may assume that $x = v_{r-1,r-1}$. Because the *first* nonzero coordinate in $\vec{d}(v_{r-1,r-1})$ is the rth coordinate where $r = d(v_{1,1}, v_{r-1,r-1})$ and the *first* nonzero coordinate in $\vec{d}(v_{r,r})$ is the $(r + 1)$th coordinate where $r + 1 = d(v_{1,1}, v_{r,r})$, it follows that $a_r = 1$ and $b_r = 0$ and so $\vec{d}(x) \neq \vec{d}(y)$.

Case 2. *x and y are both white.* First, we make some observations on the codes of vertices on B_0.

- The vertices on B_0 are the only white vertices of T whose codes contain the r-tuple $(1, 1, \ldots, 1) = 1^r$ as a subsequence. The vertex v is the only white vertex of T such that the first r coordinates of its code are 1 (that is, $\vec{d}(v) = (1^r, 0^r)$). For $1 \leq t \leq r$, the vertex $v_{0,t}$ is the only white vertex such that in $\vec{d}(v_{0,t})$ the first t coordinates and the last $r - t$ coordinates are 0 while the remaining coordinates are 1 (that is, $\vec{d}(v_{0,t}) = (0^t, 1^r, 0^{r-t})$ for $1 \leq t \leq r$). Thus, all codes of the vertices of B_0 are distinct and they are also distinct from the codes of those white vertices that are not in B_0.

Hence, we may assume that neither x nor y belongs to B_0. Let $Q_i = B_i - v = (v_{i,1}, v_{i,2}, \ldots, v_{i,r})$ be the subpath of B_i for $1 \le i \le r$. We consider two subcases, according to the location of x and y.

Subcase 2.1. $x, y \in V(Q_i)$ *where* $1 \le i \le r$. Let $x = v_{i,p}$ and $y = v_{i,q}$ where $1 \le p < q \le r$ and $p, q \ne i$.

- $\star$ First, suppose that $i \ne r$. Since (1) the *last* nonzero coordinate in $\vec{d}(v_{i,p})$ is the $(p+r)$th coordinate where $p + r = d(v_{i,p}, v_{r,r})$ and the *last* nonzero coordinate in $\vec{d}(v_{i,q})$ is the $(q+r)$th coordinate where $q + r = d(v_{i,q}, v_{r,r})$ and (2) $p < q$, it follows that $a_{q+r-1} = 0$ and $b_{q+r-1} = 1$ and so $\vec{d}(x) \ne \vec{d}(y)$.

- $\star$ Next, suppose that $i = r$. Since (1) the last nonzero coordinate in $\vec{d}(v_{r,p})$ is the $(p+r-1)$th coordinate where $p+r-1 = d(v_{i,p}, v_{r-1,r-1})$ and the last nonzero coordinate in $\vec{d}(v_{r,q})$ is the $(q+r-1)$th coordinate where $q+r-1 = d(v_{i,q}, v_{r-1,r-1})$ and (2) $p < q$, it follows that $a_{q+r-1} = 0$ and $b_{q+r-1} = 1$ and so $\vec{d}(x) \ne \vec{d}(y)$.

Subcase 2.2. $x \in V(Q_i)$ *and* $y \in V(Q_j)$ *where* $1 \le i < j \le r$. Let $x = v_{i,p}$ and $y = v_{j,q}$ where $1 \le p, q \le r$, $p \ne i$, and $q \ne j$. We consider two subcases, according to whether $p = q$ or $p \ne q$.

Subcase 2.2.1. $p = q$. Then $x = v_{i,p}$ and $y = v_{j,p}$ where $1 \le i < j \le r$ and $p \notin \{i, j\}$.

- $\star$ First, suppose that $j + 1 \le p \le r$. Since (1) the first nonzero coordinate in $\vec{d}(v_{i,p})$ is the $(p-i)$th coordinate where $p - i = d(v_{i,p}, v_{i,i})$ and the first nonzero coordinate in $\vec{d}(v_{j,p})$ is the $(p-j)$th coordinate where $p - j = d(v_{j,p}, v_{j,j})$ and (2) $i < j$, it follows that $a_{p-j} = 0$ and $b_{p-j} = 1$ and so $\vec{d}(x) \ne \vec{d}(y)$.

- $\star$ Next suppose that $1 \le p \le i - 1$. Let c_0 be the red-white coloring of T obtained by recoloring $v_{i,i}$ and $v_{j,j}$ white and all other vertices of T remain the same colors as in c. Since $d(v_{i,p}, w) = d(v_{j,p}, w)$ for every red vertex w such that $w \notin \{v_{i,i}, v_{j,j}\}$, it follows that $\vec{d}_{c_0}(x) = \vec{d}_{c_0}(y) = (f_1, f_2, \ldots, f_{2r})$. Observe that $d(v_{i,p}, v_{i,i}) = i - p$, $d(v_{i,p}, v_{j,j}) = j + p$, $d(v_{j,p}, v_{i,i}) = i + p$, and $d(v_{j,p}, v_{j,j}) = j - p$. Since $i - p < \min\{i + p, j - p, j + p\}$, it follows that $a_{i-p} = f_{i-p} + 1$ and $b_{i-p} = f_{i-p}$ and so $\vec{d}(x) \ne \vec{d}(y)$.

- $\star$ Finally, suppose that $i + 1 \le p \le j - 1$. Let c_0 be the red-white coloring of T obtained by recoloring $v_{i,i}$ and $v_{j,j}$ white and all other vertices of T remain the same colors as in c. Since $d(v_{i,p}, w) = d(v_{j,p}, w)$ for every red vertex w such that $w \notin \{v_{i,i}, v_{j,j}\}$, it follows that $\vec{d}_{c_0}(x) = \vec{d}_{c_0}(y) = (f_1, f_2, \ldots, f_{2r})$. Observe that $d(v_{i,p}, v_{i,i}) = p - i$, $d(v_{i,p}, v_{j,j}) = j + p$, $d(v_{j,p}, v_{i,i}) = i + p$, and $d(v_{j,p}, v_{j,j}) = j - p$. Since $j + p > \max\{i + p, j - p, i + p\}$, it follows that $a_{j+p} = f_{j+p} + 1$ and $b_{j+p} = f_{j+p}$ and so $\vec{d}(x) \ne \vec{d}(y)$.

Subcase 2.2.2. $p \ne q$. First, suppose that $j \ne r$. Since (1) the *last* nonzero coordinate in $\vec{d}(v_{i,p})$ is the $(p+r)$th coordinate where $p + r = d(v_{i,p}, v_{r,r})$ and the *last* nonzero coordinate in $\vec{d}(v_{j,q})$ is the $(q+r)$th coordinate where $q + r = d(v_{j,q}, v_{r,r})$ and (2) $p \ne q$, it follows that either $a_{p+r} \ne b_{p+r}$ or $a_{q+r} \ne b_{q+r}$, implying that $\vec{d}(x) \ne \vec{d}(y)$.

For simplification, we now introduce notation where a code is expressed when no 0 coordinate is given after the final nonzero coordinate of a code. For example, if a code of a vertex is a 7-tuple $(1, 0, 2, 1, 0, 0, 0)$, we simply write this code as the 4-tuple $(1, 0, 2, 1)$.

Next, suppose that $j = r$. Thus, $x = v_{i,p}$ where $1 \le i \le r - 1$ and $p \ne i$ and $y = v_{r,q}$ where $1 \le q \le r - 1$ and $p \ne q$. We consider two possibilities.

Subcase 2.2.2.1. $2 \le i \le r - 1$. First, suppose that $p \ge i + 1$. Then $\vec{d}(x) = \vec{d}(v_{i,p}) = (0^{i-p-1}, 1, 0^i, 1^{i-1}, 0, 1^{r-i})$. If $\vec{d}(y) = \vec{d}(v_{r,q})$ contains a coordinate 2, then $\vec{d}(x) \ne \vec{d}(y)$. Thus, we may assume that $d(v_{r,q}, v_{r,r}) = r - q \ne q + t$ for $1 \le t \le r - 1$ and so $\vec{d}(y) = \vec{d}(v_{r,q}) = (0^{r-q-1}, 1, 0^{2q-r}, 1^{r-1})$. Since 1^{r-1} is a subsequence in $\vec{d}(y)$ and is not a subsequence of $\vec{d}(x)$, it follows that $\vec{d}(x) \ne \vec{d}(y)$.

Next, suppose that $1 \le p \le i - 1$. If $i - p \ne p + \ell$ for some $\ell \in [r] - \{i\}$, then $(1^{i-1}, 0, 1^{r-i})$ is a subsequence of $\vec{d}(x)$ and so there is no 2 as a coordinate of $\vec{d}(x)$. If $\vec{d}(y) = \vec{d}(v_{r,q})$ contains a coordinate 2, then $\vec{d}(x) \ne \vec{d}(y)$ and so $\vec{d}(y) = \vec{d}(v_{r,q}) = (0^{r-q-1}, 1, 0^{2q-r}, 1^{r-1})$. Thus, $\vec{d}(x) \ne \vec{d}(y)$. Hence, we may assume that $i - p = p + \ell$ for some $\ell \in [r] - \{i\}$ and so $i = 2p + \ell$. This implies that there is exactly one coordinate 2 of $\vec{d}(x)$, namely $a_{i-p} = 2$. If $\vec{d}(y)$ has no coordinate 2, then $\vec{d}(x) \ne \vec{d}(y)$. Hence, we assume that $\vec{d}(y)$ has coordinate 2. This implies that $d(v_{r,q}, v_{r,r}) = r - q = q + t$ for some $t \in [r-1]$ and $b_{r-q} = 2$ is the only coordinate 2 in $\vec{d}(y)$. Hence, $i - p = r - q$ or $r - i = q - p$ and so $q > p$. There are two possibilities here. If $i - p = r - q < p + 1$, then the second nonzero coordinate in $\vec{d}(x)$ is a_{p+1} while the the second nonzero coordinate in $\vec{d}(y)$ is b_{q+1}. If $i - p = r - q > p + 1$, then the first nonzero coordinate in $\vec{d}(x)$ is a_{p+1} while the the fist nonzero coordinate in $\vec{d}(y)$ is b_{q+1}. In either case, $\vec{d}(x) \ne \vec{d}(y)$. Therefore, c is an ID-coloring and so $\mathrm{ID}(T) = r$.

Subcase 2.2.2.2. $i = 1$. Then $\vec{d}(x) = \vec{d}(v_{1,p}) = (0^{p-2}, 1, 0, 0, 1^{r-1})$. If $\vec{d}(y) = \vec{d}(v_{r,q})$ contains a coordinate 2, then $\vec{d}(x) \ne \vec{d}(y)$. Thus, we may assume that $d(v_{r,q}, v_{r,r}) = r - q \ne q + t$ for $1 \le t \le r - 1$. Since $r - q < q + r - 1$, it follows that $r - q \le q$ or $q \ge r/2$. Then $\vec{d}(y) = \vec{d}(v_{r,q}) = (0^{r-q-1}, 1, 0^{2q-r}, 1^{r-1})$. Thus, $p - 2 = r - q - 1$ (or $p + q = r + 1$), $2 = 2q - r$ (or $r = 2q - 2$ is even) and so $p = q - 1$ (or $q = p + 1$). Since r is odd, it follows that $\vec{d}(x) \ne \vec{d}(y)$. $\square$

The following is a consequence of Theorem 3.1.

Corollary 3.1. *For each odd integer $r \geq 3$, there exist a twin-free starlike tree T such that* $\mathrm{ID}(T) = r$.

In the statement of Theorem 3.1, the condition that $r \geq 3$ is an odd integer is only required in Subcase 2.2.2.2. In fact, if $r \geq 4$ is an even integer, then there are exactly two white vertices in the red-white coloring described in the proof of Theorem 3.1, namely $v_{1,p}$ and $v_{r,q}$ where $q = p + 1$ and $p + q = r + 1$, that have the same code. That is, this red-white coloring is antiregular. Therefore, we have the following.

Proposition 3.1. *For each even integer $r \geq 4$, there is an antiregular red-white coloring of the starlike tree $S_{r-1}(K_{1,r+1})$ having exactly r red vertices.*

By the technique used in the proof of Theorem 3.1, the following result can be verified.

Proposition 3.2. *Let $r \geq 4$ be an even integer. If T is the starlike tree obtained by subdividing exactly one edge of $S_{r-1}(K_{1,r+1})$, then* $\mathrm{ID}(T) = r$.

None of the trees appearing in the results just above have the identity automorphism group. We next describe a class of trees having the identity automorphism group, where each such tree is necessarily twin-free, which can be used to show that for every integer $r \geq 3$, there is a tree T with the identity automorphism group such that $\mathrm{ID}(T) = r$.

Theorem 3.2. *For each integer $r \geq 3$, there is a starlike tree T of order $1 + \binom{r+2}{2}$ having the identity automorphism group such that* $\mathrm{ID}(T) = r$.

Proof. For each integer $r \geq 3$, let $K_{1,r+1}$ be the star of order $r + 2$ with central vertex v that is adjacent to the $r + 1$ end-vertices $v_1, v_2, \ldots, v_{r+1}$. Let T be the starlike tree obtained from the star $K_{1,r+1}$ be subdividing the edge vv_i of $K_{1,r+1}$ exactly $i - 1$ times for $1 \leq i \leq r + 1$. In particular, vv_1 is not subdivided and vv_{r+1} is subdivided exactly r times. Thus, T is twin-free, the order of T is $1 + \binom{r+2}{2}$ and $\mathrm{diam}(T) = 2r + 1$. Since no two vertices of T are similar, it follows that T has the identity automorphism group. For each integer i with $1 \leq i \leq r + 1$, let $B_i = (v, v_{i,1}, v_{i,2}, \ldots, v_{i,i})$ be a branch of T at v.

First, we show that $\mathrm{ID}(T) \geq r$. For any red-white coloring of T that assigns the color red to at most $r - 1$ vertices of T, there are at least two branches B_i and B_j of T at v such that the paths $B_i - v$ and $B_j - v$ contain no red vertices of T. However then, $\vec{d}(v_{i,1}) = \vec{d}(v_{j,1})$, for example, and so this red-white coloring is not an ID-coloring of T. Therefore, $\mathrm{ID}(T) \geq r$. Next, we show that T has an ID-coloring with exactly r red vertices. Define a red-white coloring c of T by assigning the color red to each vertex $v_{i,i}$ for $1 \leq i \leq r$ and white to the remaining vertices of T. Thus, T has exactly r red vertices. It remains to show that c is an ID-coloring of T. Since $\mathrm{diam}(T) = 2r + 1$, the code of each vertex of T is a $(2r + 1)$-vector. Let x and y be two distinct vertices of T. We consider two cases, according to whether x and y are both red or both white. Let $\vec{d}(x) = (a_1, a_2, \ldots, a_{2r+1})$ and $\vec{d}(y) = (b_1, b_2, \ldots, b_{2r+1})$.

Case 1. *x and y are both red*. Let $x = v_{i,i}$ and $y = v_{j,j}$ where $1 \leq i < j \leq r$.

- ⋆ First, suppose that $j \neq r$. Since (1) the *last* nonzero coordinate in $\vec{d}(v_{i,i})$ is the $(i + r)$th coordinate where $i + r = d(v_{i,i}, v_{r,r})$ and the *last* nonzero coordinate in $\vec{d}(v_{j,j})$ is the $(j + r)$th coordinate where $j + r = d(v_{j,j}, v_{r,r})$ and (2) $i < j$, it follows that $a_{j+r} = 0$ and $b_{j+r} = 1$ and so $\vec{d}(x) \neq \vec{d}(y)$.
- ⋆ Next, suppose that $j = r$. We saw that the last nonzero coordinate in $\vec{d}(v_{i,i})$ where $1 \leq i \leq r - 1$ is the $(i + r)$th coordinate. Since the last nonzero coordinate in $\vec{d}(v_{r,r})$ is the $(2r - 1)$th coordinate where $2r - 1 = d(v_{r-1,r-1}, v_{r,r})$, it follows that if $i \neq r - 1$, then $\vec{d}(x) \neq \vec{d}(y)$. Thus, we may assume that $x = v_{r-1,r-1}$. Because the *first* nonzero coordinate in $\vec{d}(v_{r-1,r-1})$ is the rth coordinate where $r = d(v_{1,1}, v_{r-1,r-1})$ and the *first* nonzero coordinate in $\vec{d}(v_{r,r})$ is the $(r + 1)$th coordinate where $r + 1 = d(v_{1,1}, v_{r,r})$, it follows that $a_r = 1$ and $b_r = 0$ and so $\vec{d}(x) \neq \vec{d}(y)$.

Case 2. *x and y are both white*. First, we verify the following claim.

Claim. If $x \in V(B_{r+1})$ or $y \in V(B_{r+1})$, then $\vec{d}(x) \neq \vec{d}(y)$.

The vertices on B_{r+1} are the only white vertices of T whose codes contain the r-tuple $(1, 1, \ldots, 1) = 1^r$ as a subsequence. The vertex v is the only white vertex of T such that the first r coordinates of its code are 1 (that is, $\vec{d}(v) = (1^r, 0^{r+1})$). For $1 \leq t \leq r + 1$, the vertex $v_{r+1,t}$ is the only white vertex such that in $\vec{d}(v_{r+1,t})$ the first t coordinates and the last $r + 1 - t$ coordinates are 0 while the remaining coordinates are 1 (that is, $\vec{d}(v_{r+1,t}) = (0^t, 1^r, 0^{r+1-t})$ for $1 \leq t \leq r$). Thus, all codes of the vertices of B_{r+1} are distinct and they are also distinct from the codes of those white vertices that are not in B_{r+1}. Hence, the claim holds.

By the claim, we may assume that $x \notin V(B_{r+1})$ and $y \notin V(B_{r+1})$. Let $Q_i = B_i - v = (v_{i,1}, v_{i,2}, \ldots, v_{i,i})$ be the subpath of B_i for $2 \leq i \leq r$. We consider two subcases, according to the location of x and y.

Subcase 2.1. *$x, y \in V(Q_i)$ where $2 \leq i \leq r$*. Let $x = v_{i,p}$ and $y = v_{i,q}$ where $1 \leq p < q \leq i - 1$.

- ⋆ First, suppose that $i \neq r$. Since (1) the *last* nonzero coordinate in $\vec{d}(v_{i,p})$ is the $(p+r)$th coordinate where $p+r = d(v_{i,p}, v_{r,r})$ and the *last* nonzero coordinate in $\vec{d}(v_{i,q})$ is the $(q+r)$th coordinate where $q+r = d(v_{i,q}, v_{r,r})$ and (2) $p < q$, it follows that $a_{q+r-1} = 0$ and $b_{q+r-1} = 1$ and so $\vec{d}(x) \neq \vec{d}(y)$.

- ⋆ Next, suppose that $i = r$. Since (1) the last nonzero coordinate in $\vec{d}(v_{r,p})$ is the $(p+r-1)$th coordinate where $p+r-1 = d(v_{i,p}, v_{r-1,r-1})$ and the last nonzero coordinate in $\vec{d}(v_{r,q})$ is the $(q+r-1)$th coordinate where $q+r-1 = d(v_{i,q}, v_{r-1,r-1})$ and (2) $p < q$, it follows that $a_{q+r-1} = 0$ and $b_{q+r-1} = 1$ and so $\vec{d}(x) \neq \vec{d}(y)$.

Subcase 2.2. $x \in V(Q_i)$ *and* $y \in V(Q_j)$ *where* $2 \leq i < j \leq r$. Let $x = v_{i,p}$ and $y = v_{j,q}$ where $1 \leq p \leq i-1$ and $1 \leq q \leq j-1$. We consider two subcases, according to whether $p = q$ or $p \neq q$.

Subcase 2.2.1. $p = q$. Then $x = v_{i,p}$ and $y = v_{j,p}$ where $2 \leq i < j \leq r$ and $p \notin \{i, j\}$. Let c_0 be the red-white coloring of T obtained by recoloring $v_{i,i}$ and $v_{j,j}$ white and all other vertices of T remain the same colors as in c. Since $d(v_{i,p}, w) = d(v_{j,p}, w)$ for every red vertex w such that $w \notin \{v_{i,i}, v_{j,j}\}$, it follows that $\vec{d}_{c_0}(x) = \vec{d}_{c_0}(y) = (f_1, f_2, \ldots, f_{2r})$. Observe that $d(v_{i,p}, v_{i,i}) = i - p$, $d(v_{i,p}, v_{j,j}) = j + p$, $d(v_{j,p}, v_{i,i}) = i + p$, and $d(v_{j,p}, v_{j,j}) = j - p$. Since $i - p < \min\{i+p, j-p, j+p\}$, it follows that $a_{i-p} = f_{i-p} + 1$ and $b_{i-p} = f_{i-p}$ and so $\vec{d}(x) \neq \vec{d}(y)$.

Subcase 2.2.2. $p \neq q$. First, suppose that $j \neq r$. Since (1) the *last* nonzero coordinate in $\vec{d}(v_{i,p})$ is the $(p+r)$th coordinate where $p+r = d(v_{i,p}, v_{r,r})$ and the *last* nonzero coordinate in $\vec{d}(v_{j,q})$ is the $(q+r)$th coordinate where $q+r = d(v_{j,q}, v_{r,r})$ and (2) $p \neq q$, it follows that either $a_{p+r} \neq b_{p+r}$ or $a_{q+r} \neq b_{q+r}$, implying that $\vec{d}(x) \neq \vec{d}(y)$.

Next, suppose that $j = r$. Thus, $x = v_{i,p}$ where $2 \leq i \leq r-1$ and $1 \leq p \leq i-1$ and $y = v_{r,q}$ where $1 \leq q \leq r-1$ and $p \neq q$. If $i - p \neq p + \ell$ for some $\ell \in [r] - \{i\}$, then $(1^{i-1}, 0, 1^{r-i})$ is a subsequence of $\vec{d}(x)$ and so there is no coordinate 2 in $\vec{d}(x)$. If $\vec{d}(y) = \vec{d}(v_{r,q})$ contains 2 as a coordinate, then $\vec{d}(x) \neq \vec{d}(y)$ and so $\vec{d}(y) = \vec{d}(v_{r,q}) = (0^{r-q-1}, 1, 0^{2q-r}, 1^{r-1}, 0, \ldots, 0)$. Thus, $\vec{d}(x) \neq \vec{d}(y)$. Hence, we may assume that $i - p = p + \ell$ for some $\ell \in [r] - \{i\}$ and so $i = 2p + \ell$. This implies that there is exactly one coordinate of $\vec{d}(x)$ which is 2, namely $a_{i-p} = 2$. If $\vec{d}(y)$ has no coordinate 2, then $\vec{d}(x) \neq \vec{d}(y)$. Hence, we assume that $\vec{d}(y)$ has 2 as a coordinate. This implies that $d(v_{r,q}, v_{r,r}) = r - q = q + t$ for some $t \in [r-1]$ and $b_{r-q} = 2$ is the only coordinate 2 in $\vec{d}(y)$. Hence, $i - p = r - q$ or $r - i = q - p$ and so $q > p$. There are two possibilities here. If $i - p = r - q < p + 1$, then the second nonzero coordinate in $\vec{d}(x)$ is a_{p+1} while the the second nonzero coordinate in $\vec{d}(y)$ is b_{q+1}. If $i - p = r - q > p + 1$, then the first nonzero coordinate in $\vec{d}(x)$ is a_{p+1} while the the fist nonzero coordinate in $\vec{d}(y)$ is b_{q+1}. In either case, $\vec{d}(x) \neq \vec{d}(y)$.

Therefore, c is an ID-coloring and so $\text{ID}(T) = r$. □

Several problems are suggested by the results presented here.

(1) For a given integer $r \geq 3$, what is the smallest order of a tree T such that $\text{ID}(T) = r$?

(2) For a given integer $r \geq 3$, what is the smallest order of a twin-free tree T such that $\text{ID}(T) = r$?

For (2), we have seen that this smallest order is no more than $1 + \binom{r+2}{2}$.

Acknowledgment

We are grateful to Prof. Gary Chartrand for suggesting us concepts and kindly providing useful information on this topic.

References

[1] A. Ali, A survey of antiregular graphs, *Contrib. Math.* **1** (2020) 67–79.
[2] A. Ali, G. Chartrand, P. Zhang, *Irregularity in Graphs*, Springer, New York, 2021.
[3] G. G. Chappell, J. Gimbel, C. Hartman, Bounds on the metric and partition dimensions of a graph, *Ars Combin.* **88** (2008) 349–366.
[4] G. Chartrand, L. Eroh, M. A. Johnson, O. R. Oellermann, Resolvability in graphs and the metric dimension of a graph, *Discrete Appl. Math.* **105** (2000) 99–113.
[5] G. Chartrand, Y. Kono, P. Zhang, Distance vertex identification in graphs, *J. Interconnection Netw.* **21** (2021) Art# 2150005.
[6] G. Chartrand, E. Salehi, P. Zhang, The partition dimension of a graph, *Aequationes Math.* **59** (2000) 45–54.
[7] G. Chartrand, P. Zhang, The theory and applications of resolvability in graphs – a survey, *Congr. Numer.* **160** (2003) 47–68.
[8] F. Foucaud, G. B. Mertzios, R. Naserasr, A. Parreau, P. Valicov, Identification, location-domination and metric dimension on interval and permutation graphs. I. Bounds, *Theoret. Comput. Sci.* **68** (2017) 43–58.
[9] F. Foucaud, G. B. Mertzios, R. Naserasr, A. Parreau, P. Valicov, Identification, location-domination and metric dimension on interval and permutation graphs. II. Algorithms and complexity, *Algorithmica* **78** (2017) 914–944.
[10] F. Harary, R. A. Melter, On the metric dimension of a graph, *Ars Combin.* **2** (1976) 191–195.
[11] M. Hauptmann, R. Schmied, C. Viehmann, Approximation complexity of metric dimension problem, *J. Discrete Algorithms* **14** (2012) 214–222.
[12] C. Hernando, M. Mora, M. I. Pelayo, C. Seara, D. Wood, R. David, Extremal graph theory for metric dimension and diameter, *Electron. J. Combin.* **17** (2010) Art# R30.
[13] M. A. Johnson, Structure-activity maps for visualizing the graph variables arising in drug design, *J. Biopharm. Statist.* **3** (1993) 203–236.
[14] M. A. Johnson, Browsable structure-activity datasets, In: R. Carbó-Dorca, P. G. Mezey (Eds.), *Advances in Molecular Similarity*, Vol. 2, JAI Press Inc., Stamford, 1998, pp. 153–170.
[15] P. J. Slater, Leaves of trees, *Congr. Numer.* **14** (1975) 549–559.
[16] P. J. Slater, Dominating and reference sets in graphs, *J. Math. Phys. Sci.* **22** (1988) 445–455.
[17] P. Zhang, *Color-Induced Graph Colorings*, Springer, New York, 2015.
[18] P. Zhang, *A Kaleidoscopic View of Graph Colorings*, Springer, New York, 2016.